Toward Interactive Audiovisual Cartography: Motivations, Design Strategies, and Methods

by

H. Glenn Brauen

A thesis submitted to the Faculty of Graduate and Postdoctoral Affairs in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Geography

Carleton University
Ottawa, Ontario

© 2011

H. Glenn Brauen

Library and Archives Canada

Published Heritage Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et Archives Canada

Direction du Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-87771-5

Our file Notre référence
ISBN: 978-0-494-87771-5

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distrbute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protege cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Abstract

Maps are still predominantly understood as silent. This thesis argues that sound can play a more significant role within a reconceptualized study of cartography, introduced here using the term 'audiovisual cartography,' in which visual and acoustic design complement each other and together provide alternative possibilities for assisting in the examination and communication of complex spatial information concerning a wide variety of subject matters. This thesis examines theoretical, design, technological, and user acceptance and comprehension challenges for audiovisual cartography, in particular, interactive audiovisual maps designed for use and distribution over digital networks. Drawing on previous research literatures concerning sound design for films and games, and research concerning predominantly silent interactive and animated maps, a theoretical framework for audiovisual cartography and tentative guidance for a variety of ways in which sound may be used to create, add, or reinforce meaning as part of maps are proposed. A model for implementing interactive audiovisual map applications is outlined and a portable, cross-platform software implementation meeting the requirements of this model is introduced. The results of user evaluations with a set of interactive audiovisual map prototypes are discussed. Finally, an example application, *Canada-USA Commodity Trade 1976–2000*, demonstrating a set of design ideas and highlighting issues concerning the development, distribution, and use of interactive and animated audiovisual maps is discussed. This research contributes new

ideas for the design of interactive maps and visualization tools for spatial data, and provides an open-source sound subsystem implementation upon which additional research could be based.

Acknowledgements

This thesis would not have taken on the shape it has without the people with whom I connected while working on it. I thank the members of my advisory committee, Dr. D. R. Fraser Taylor, Dr. Paul Théberge, and Dr. Iain Wallace, each of whom provided useful guidance and constructive criticisms throughout this research and my program at Carleton University.

Dr. Taylor, as director of Carleton University's Geomatics and Cartographic Research Centre (GCRC), was instrumental in providing research facilities, establishing a transdisciplinary research team and collaborative networks tackling a range of projects, and financial support within which I was able to position, develop, and extend the research project reported here. Barbara George has provided administrative and financial management for all of these activities.

Early in my studies, a group of researchers interested in the use of sound as a cartographic modality collaborated under the auspices of the Cybercartography and the New Economy project, funded by the Social Sciences and Humanities Research Council (SSHRC) of Canada under the Initiative on the New Economy (INE) Collaborative Research Initiative Grant for which Dr. Taylor was the principal investigator. The 'sound cluster', as it came to be known, kindled my interest in pursuing the theory of film sound as one source of guidance for interactive audiovisual cartography. I thank Dr. Paul Théberge, Dr. Sébastien Caquard, Benjamin Wright, and Paul Jasen for their inspiration and collaborations that led to, among other outcomes, map examples used in evaluation sessions discussed in this thesis.

This research has benefitted from the prior work and assistance of students and researchers at the GCRC. Peter Pulsifer organized research and the development of the *Cybercartographic Atlas of Antarctica* on which I and the sound cluster collaborated and for which two map examples discussed in this thesis were created. Brian Eddy organized research and the development of the *Cybercartographic Atlas of Canada's Trade with the World* which provided data sets and research upon which I built when creating map examples for this thesis. Amos Hayes and Jean-Pierre Fiset guided the development of the encompassing *Nunaliit Cybercartographic Atlas Software Framework* within which early versions of the cartographic sound subsystem discussed in this thesis were created and released, and have continued to collaborate on software infrastructures needed for interactive web maps like those discussed here. Special thanks to Tracey Lauriault for her assistance during the arduous transcription and checking of user evaluation session videos for this research, in addition to her critical thinking and constructive criticisms on various interim prototypes. Many of these collaborations within the GCRC, and others not directly related to this thesis, have turned into friendships and I thank you all.

The user evaluations conducted for and discussed as part of this thesis were planned and designed with the assistance of Dr. Avi Parush and Dr. Sébastien Caquard. The user evaluation sessions were conducted by Shelley Roberts and Aren Hunter, under Dr. Parush's supervision, and in collaboration with the author. Funding to conduct the usability study was provided by SSHRC under an Image, Text, Sound, and Technology Research Grant for which Dr. D. R. Fraser Taylor was the principal investigator.

My research has been funded, in part, through a SSHRC Canada Graduate Scholar Doctoral Fellowship and through an Ontario Graduate Scholarship.

Finally, to Lydia Asmus, a light in my heart. Thank you for your patience throughout the writing of this thesis and especially with my tendency to lose track of time while working.

Contents

List of Tables

List of Figures

List of Appendices

Part I

Audiovisual Maps: Introduction, Objectives and Context

Geographic and cartographic research concerning sound is not without precedent. But as will be outlined in Chapter 1, the research literature focused on conceptions of sound in geography is not extensive, and even less has been written with respect to sound in cartography. The research discussed in this thesis examines possibilities for incorporating sound as a component of audiovisual map designs distributed and used digitally. From the outset, several questions concerning the practical usefulness of such a proposal arise. Why is sound of use as a component of maps? How can sound be distributed and controlled as part of map artefacts? How can map designs incorporating sound be created? How will users of such maps accept, understand, and use them?

Despite the continued applicability and popularity of paper maps for certain purposes, maps increasingly are being used as digital media and are being distributed and used over digital networks. Although the introduction of dynamic map products has changed and continues to change the material form of maps (for *some* users), the adoption of sound as a component of these, has been slow.[1] This

[1]Dynamic map products and processes have been researched under a variety of names including *interactive maps* (Asche and Herrmann 1994; Dorling 1992; Peterson 1995), *animated cartography* (Campbell and Egbert 1990; Harrower 2003, 2007; Kraak and Klomp 1995; Peterson 1995; Thrower 1959), *multimedia cartography* (Cartwright and Peterson 2007), *cybercartography* (Taylor 2003, 2005c; Taylor and Pyne 2010), *web mapping* (Kraak and Brown 2000) and *maps for the Internet* (Peterson 2003).

despite the growing predominance of digital media distribution, with and without the inclusion of sound, over the same means of distribution and under a variety of economic models. Network bandwidths have grown such that for many consumers digital audio is predominantly acquired through digital downloading.

The growing predominance of acoustic media distribution over digital networks and the increasing predominance of digital map use appear to be trends that are operating in isolation from each other. A recent statement of research agendas by the International Cartographic Association (Virrantaus et al. 2009) discussed the use of sound in conjunction with visual maps only with respect to usability research for visually impaired users.

The central argument of this thesis is that sound can play a more significant role within a reconceptualized study of cartography, introduced here using the term 'audiovisual cartography,' in which visual and acoustic design complement each other and together provide alternative possibilities for assisting in the examination and communication of complex spatial information concerning a wide variety of subject matters. Theoretical, technological, and design challenges must be examined and overcome for audiovisual cartography to be realized as a practical method for such purposes. The objective of this thesis is to examine the possibilities for expanding the consideration of sound as a component of digital maps distributed and used over networks such as the World Wide Web. To that end, three parallel, inter-related research threads are pursued within this dissertation:

- **Motivations**: In what ways could sound complement a predominantly visual digital map? What existing research into the use of sound in other audiovisual media could provide guidance for the use of sound as part of an audiovisual cartography? Turning to previous cartographic research, what purposes could sound serve in studying problems which are being addressed by geographers and cartographers using other means? Are there analogies to be drawn from research into visual interactive and animated cartography that could provide tentative working models for audiovisual cartography?

- **Design Strategies**: In some ways, this is the most nebulous of the research threads addressed in this thesis. Broad design guidelines may often be derived to apply to categories of projects but, at the outset, designs for audiovisual cartography must be developed to address requirements of specific projects and the users of those projects. If specific audiovisual map examples are developed by applying research into motivations for the use of sound in audiovisual mapping, are they effective with respect to project objectives and how do users respond to them? Are users able to work with the examples and do they understand the information being presented? What type of guidance or explanation do they need to work with the audiovisual designs? Do they accept the use of sound as part of the designs? Can tentative design guidelines be enunciated?
- **Methods**: What are the obstacles to distributing and using sound as part of audiovisual map designs distributed over networks such as the World Wide Web? What types of audio should a system for managing sound as part of such designs support and what capabilities are required to present sound as part of a combined audiovisual map?

This thesis examines each of these concerns. Part I sets the context for this research by first arguing that conceptions of human sensory perception have varied through time and in different cultures and, although historically there has been a visual bias within geographical and cartographic research, recent geographical research initiatives appear to be more open to research examining other modalities, including aurality. Whereas cartographic research with its predominantly paper-based history has had an understandable visual bias, trends in digital technologies during the past several decades have created economic, social, and technological conditions in which other modalities may be more fully supported, with listening being the most obvious possibility for increased consideration. Next, consideration is given to the growing predominance of digital media usage supported and stimulated by telecommunications networks and, in particular, the phenomena of concurrent use by individuals of multiple media to suggest that despite the bias toward the design of silent maps, *silent map use* may be less common than is supposed. Lastly, audiovisual cartography is discussed theoretically, based on the conception of interactive, animated, and audiovisual maps as digital media, and a new framework for discussing acoustic signification within audiovisual cartography is presented.

Part II outlines a range of possible motivations for using sound in audiovisual maps, drawing on literatures concerning other media that combine visual and acoustic modalities: film, games, and human-computer interaction research. This discussion of how sound has been used in other media is augmented through an examination of research into interactive and animated cartography, mapping forms that explicitly use time in the representation of spatial information, to argue that sound can be used in analogical ways in an audiovisual cartography to address similar mapping problems.

Part III builds on earlier sections of the thesis by examining questions concerning how users respond to audiovisual map applications and how audiovisual map designers could design and distribute audiovisual maps over the World Wide Web. Proceeding approximately chronologically with respect to the order in which this research was conducted, a study concerning user evaluations of a set of audiovisual map applications is first presented and the responses of the users to these maps are discussed. Based on lessons from those evaluations and experience gained through preparing (collaborating to prepare, in some cases) the audiovisual map examples used in those evaluations, an overview of a software subsystem redesigned to support the design and distribution of audiovisual maps over the World Wide Web is presented. The high-level design of this software subsystem extends the theoretical framework presented earlier in the thesis by establishing and explaining a model through which a sound design can be specified for a particular audiovisual map application. An example application, *Canada-USA Commodity Trade 1976-2000*, created concurrently with the sound subsystem redesign, both as a case study to examine design issues for audiovisual maps and as a test case used to exercise the redesigned sound subsystem, is discussed. Finally, outcomes from this research and possible future research directions are discussed.

The research described here was conceived in the context of cybercartographic atlas projects. Taylor (2003, 406) defined *cybercartography* as "the organization, presentation, analysis and communication of spatially referenced information on a wide variety of topics of interest and use to society in an interactive, dynamic, multimedia, multisensory and multidisciplinary format." Among the (seven) major elements of cybercartography outlined by Taylor (2003, 2005c), the following are particularly relevant to this discussion: cybercartography is multisensory and not strictly visual; cybercartography uses a range of media forms and distributes them for use via telecommunication technologies such as the World Wide Web; and cybercartography adapts interactive media to engage and inform users.

Although the research described here developed in the context of cybercartographic atlas projects and the use of sound as a representational modality for spatialized information follows directly from the multisensory and multimodal emphasis of cybercartography, the use of sound described in this dissertation is an appropriate extension to other conceptions of contemporary cartography such as multimedia cartography (Cartwright et al. 2007), maps for Internet distribution (Peterson 2003, 2008), and interactive cartography (Asche and Herrmann 1994; Dorling 1992; Peterson 1995). For that reason, I have chosen to use the term *audiovisual cartography* to describe this research.

The International Cartographic Association (ICA), at its 10th General Assembly in 1995, adopted the following statement as the definition of a *map*:

> A map is a symbolised image of geographical reality, representing selected features or characteristics, resulting from the creative effort of its author's execution of choices, and is designed for use when spatial relationships are of primary relevance.[2]

By contrast, Harley and Woodward (1987, xvi) offered a definition that allowed a broader range of topics to be examined using maps by including "things, concepts, conditions, processes, or events in the human world." Harmon (2003, 11) extended the subject matter of maps into the realms of the imagination "beyond the boundaries of geography or convention." For the purposes of examining the design possibilities of an audiovisual cartography, any of these definitions would be a reasonable basis if the modal restriction requiring only graphic or image-based representation of the spatial subject matter was loosened. To simplify matters in this thesis, the visual components of the map examples to be discussed will be confined to two-dimensional orthographic map views and will not include bird's eye view depictions or three-dimensional representations. Although sound could certainly be used along with these other types of graphic representations, including interactive and animated versions of these representations in this examination would create definitional difficulties by further blurring the distinctions across a

[2]This statement is currently incorporated into the ICA mission statement: http://icaci.org/mission. Last modified 2008. Accessed July 2011.

range of media including maps, three-dimensional models and fly-throughs, multimedia, and film.

The discussion in this thesis is limited to the consideration of computer-based representations incorporating prerecorded or composed sounds used in coordination with visual elements consisting of text, images, and graphical elements, and more specifically audiovisual representations of spatial phenomena. These types of representations are suitable for use with either stand-alone or network-distributed maps and atlases and can also be seen as an extension of a type of representation examined by earlier literature on animated and interactive cartography which may have, but generally did not, include the use of sound.

1

Listening: Perception and Disciplinary Considerations

Sound is typically not considered in methodological discussions of cartography and is only cursorily considered if at all, with very few exceptions, in discussions of the design and use of computer-based cartographic and geographic information systems. This chapter discusses changing understandings of sensory perceptions and considerations of sound and aurality within the disciplines of geography and cartography. First, a brief survey of literature is presented to argue that understandings of human sensory perception and the priority of each sense are learned and culture-based rather than 'natural.' Then, literature surveys are presented highlighting the visual biases in research concerning geography and cartography but also outlining the extent to which sound and aurality have emerged as subjects in their own right and as theoretical perspectives, especially in recent decades.

1.1 Sensory Perceptions

The human sensory system is a fundamental means by which we gain knowledge of our environment and interact with that environment. Within Western philosophy, the understanding of the role played by the human senses in our perceptions of the world has varied according to the ontologies and epistemologies of different cultures and groups within society. Understandings of the number, enumeration, and ordering of the senses have varied, as have understandings of relations between the senses and faculties of human thought. Aristotle (1931a, b) enumerated five senses: sight, hearing, smell, touch, and taste. This classification is convenient in creating a correlation between sensory modes and discrete sense organs but alternative classifications have been proposed. Gold (1980, 50) listed ten senses whereas Downs and Stea (1977, 23) discuss perception as "a synthesis of [only four] different types of information: visual, auditory, olfactory, and kinaesthetic."

Western philosophy has generally maintained vision as preeminent among the senses (Classen 2005; Synnott 1991), although it is questionable whether visual practice has been as unified or homogeneous as such statements concerning its dominance seem to suggest (Jay 1988). Set against the primacy of vision, other sensory modes have been ranked as less important, more susceptible to temptation, and gendered (Classen 2005). The intimate senses of touch, taste, and smell were during the Renaissance thought to be feminine, "nurturing, seductive, dissolute in [their] merging of self and other", whereas sight and hearing were predominantly understood as masculine, "dominating, rational, orderly" (ibid., 70). But against this ranking of both hearing and vision as more valued than other modalities, Kozloff (2000, 9–14) argued that speech has received differentiated consideration based on gendered understandings of speech contexts. "Whenever speech is valued as an important act in a public sphere, it is seen as masculine; when it is held to no account in the casual language of ordinary conversation, it is ascribed to women" (ibid., 11). This contextual acceptance of the importance of speech at times has inverted the ranking of vision and listening. According to Soskice (1996, 31), although vision was considered as primary among the senses during the Middle Ages, it was not in all contexts accepted as providing access to knowledge:

> The medievals [...] privileged vision above the other senses. But this primacy of vision was qualified, particularly in matters epistemological, by at least two important reservations. The first was the accepted inadequacy of all the senses, including vision, to obtain the most important kind of knowledge, viz, knowledge of God; and the second is the importance accorded to "the Word."

The senses have been the subject of scientific, philosophical, and theological debate. Although Saint Ignatius Loyola had been prepared to believe that "the white object [he] saw was black, if that was the wisdom of the church" (Synnott 1991, 71), scholars such as Descartes and Locke were more confident in rationality as a means to understand the world. Descartes, inspired by mathematical rationalism and distrusting sensory perception, placed reason above the senses. Locke believed the senses were the foundation of thought and nothing could be known without it first being sensed (Classen 1993, 4). Contradicting his own philosophical distrust of the senses in his scientific writings, Descartes argued that all actions in life relied upon sensory perception and, singling out sight, stated that the greatest human inventions are those that augment the power of sight (Pickles 2004, 81; Synnott 1991, 69–71).

The list of technologies that may have enhanced human vision is long. McLuhan (1962) argued that writing fundamentally altered the human perception of space, leading to the dominance of vision. Roberts (2005) argued that, as scientific techniques in chemistry developed, precise instruments began to replace the use of direct sensory information as part of investigative procedure. Increased use of instrumental procedure in chemistry was the result of profound and much debated epistemological changes that redefined the nature of evidence, methodology and the types of problems the science would investigate. "A major component of the late eighteenth century chemical revolution entailed the introduction of investigative procedures that subordinated human sense to the calibrated readings of highly complex experimental apparatus" (ibid., 106). The importance of the use of touch, taste, smell, and hearing were all reduced while the importance of sight increased with respect to reading the results produced by laboratory instruments but decreased with respect to direct judgments of the physical characteristics of material (ibid., 123).

As with uncertainties in orderings of the senses outlined by Soskice, practice and valuations of sensory perceptions may be understood as highly contextual. Sterne (2003) argued that historical development of acoustics as a distinct discipline was dependent on innovations that enabled the visual study of sound. Understandings of sound, as a general phenomenon caused by vibrational energy transmitted as waves, were achieved through visual means such as patterns created in sand subjected to acoustic vibrations (ibid., 44) and visual tracings of sound waves created by attaching a stylus to a vibrating membrane, modelled after or indeed using a tympanum (ibid., 46–51). But from the late eighteenth century and continuing into the twentieth century, "[l]istening gets articulated to notions of science, reason, and rationality [... and ...] becomes a technical skill, a skill that can be developed and used toward instrumental ends" (ibid., 93). Sterne argued this was so in the disciplinary practice of physicians, symbolized by their use of and association with the stethoscope, in the practice of telegraphers and, later, in the growing middle class's adoption and use of sound reproduction technologies.

Despite the status of differentiated sensory practice within certain disciplines or cultures, other senses continue to be used in everyday experience of the world. Within certain non-Western cultures, vision has not been predominant to the same extent. Carpenter (1973) described travelling by kayak with Aivilik hunters and noted their abilities to navigate in conditions in which he could discern no outline of landscape, instead relying on the sounds of surf and nesting birds, the smells of the shore, and the feel of wind, spray, and wave patterns. He stated that when "they used their eyes, it was often with an acuity that amazed [him]" (ibid., 36) but they weren't lost when visibility was poor.

According to Tuan (1974, 12), young children learn about space, movement, and body coordination through unstructured play, and with close contact to people they learn to communicate, whereupon "how their capacities are used and develop begin to diverge." Cultural attitudes toward the use and priority given to the different senses accompanied by different levels of training of each sense, based on those attitudes, lead to different experiences of the world. Howes (1991a, 8, emphasis in the original) argued that "it becomes possible to think of cultures as contrasting in terms of the distinctive patterns to the *interplay of the senses* they present." Within cultures, differences in sensory experience can be observed between age groups, genders, and socio-economic classes (Rodaway 1994, 22). How each sense is stimulated, the presence of stimuli within the environment, the manner of transmission of those stimuli, and how attuned an individual is to them all affect the way in which a particular sense can structure the world for an individual.

A rebalancing of the senses can also be caused by perceptual deficiency such as blindness or visual impairment. Sound fills an important role in defining the world for blind people. Whereas sight creates the world as a set scene, a set of objects available to the glance, "the sentient participates in the auditory world and it unfolds over time and continually changes" (Rodaway 1994, 102). Hull, who became blind as an adult, described the active character of what he considered to be good weather:

> The idea of a nice day is largely visual. A nice day occurs when there is a clear blue sky [...] For me, the wind has taken the place of the sun, and a nice day is a day when there is a mild breeze. This brings into life all the sounds in my environment. The leaves are rustling, bits of paper are blowing along the pavement, the walls and corners of large buildings stand out under the impact of the wind, which I feel in my hair and on my face, in my clothes (Hull 1990, 12, quoted in Rodaway 1994, 103).

Thus not only are the environmental cues available to an individual dependent on the senses through which they engage the world but the subjective emotional assessment and the delineation of experiences of that world created through the mediation of the senses is also dependent on which senses are available or on the priority placed on each of the available senses.

1.2 Geography: Vision and Aurality

> Geographers appear to require that the demonstration of their theories be supplied by the visual evidence of the world around them. Theirs finally is the argument of the eye. (Cosgrove 1998, 31)

Vision has been widely privileged within geographic research as the primary means of understanding the complexities of the environment and the role of humans within that environment. Pocock (1989) argued that the neglect of sound as a subject and as a method of inquiry was surprising because of the important role sound plays in interactions with our environment in everyday life but was to be expected because of the primacy of our visual sense, at least in western culture. Rodaway (1994, 35) argued "it is important to remember that [the study of senses in isolation] is an abstraction. [...] all geographical experiences are made up of a complex of sensuous information combining activities of the sense organs, the body and its limbs, and mental processes."

Haggett (1965, 2), outlining methods appropriate to geography conceived as a science based on locational and distributional analysis, declared the eye to be "a necessary part of our scientific equipment in that pattern and order exist in knowing what to look for, and how to look." The visual map has been widely accepted within the discipline as both an important analytical tool and as a succinct means of presenting conclusions (Thrower 1972). O'Sullivan and Unwin (2003, 17, 20) also noted the frequency with which map-based displays are used, both as aids in the generation of ideas concerning spatial patterns and to present evidence, but cautioned that there is a need to apply other techniques to validate "the significance or importance of the apparently obvious." Visual imagery, in various forms, has been privileged as the primary means of presenting evidence to support geographical research in seminars, conferences, and classrooms (Crang 2003; Rose 2003, cf. Matless 2003, cf. Driver 2003). Despite their focus on the visual elements of geographical presentations, both Rose (2003) and Matless (2003) note the presence of a lecturer addressing an audience. Sound, especially speech, plays an important role in these settings.

Cosgrove (1998) argued that *landscape* is a fundamental concept in geographical research and furthermore that it is an inherently visual concept. Cosgrove argued that landscape has two distinct but related meanings: an artistic perspective on scenery from a specific viewpoint; and "the integration of natural and human phenomena which can be empirically verified and analysed by the methods of scientific enquiry over a delimited portion of the earth's surface" (ibid., 9). Cosgrove (1998, 28–31) argued that the concept of landscape, including a predominantly visual bias in both the application of the second of the above definitions and in its disciplinary scientific roots, underpinned a diversity of geographical methods and tools ranging from regional description to nomothetic, theoretical methodologies.

Because of the widespread development of visual technologies (e.g., linear perspective, cartography) and the deployment of those technologies by political interests, Western visuality has been conceptualized as intertwining with and supporting material projects of control, exploitation, and domination (Cosgrove 1998; Gregory 2001). The distancing of the observer from the subject of study through the practiced gaze has been theorized as gendered activity in which dominating, masculine, scientific inquiry is focused on feminized nature (Rose 1996). But as Kwan (2002) argued in her appraisal of Geographic Information Systems (GIS), vision is neither inherently bad nor good. The social effects of the use of GIS are more important than the modalities through which research is conducted. Kwan argued that the incorporation of qualitative data, in the form of digitized voice clips, videos, hand-drawn sketches and maps, and handwriting can balance the synoptic viewpoint of maps, can be used to construct spatial narratives, and to compile biographical sketches.

The understanding of the human sensory order as necessarily dominated by vision is not universally accepted and the importance of aural epistemologies have been shown to recast understandings of social spaces. Schafer (1977) argued that the sound environment within which societies live is an indication of the power relationships in those societies. What noises exist and what are the impacts of those sounds on the inhabitants? Who controls the loudest sounds and what is the response of the rest of the community to those sounds? Following Schafer's concept of the *soundscape* (ibid., 7–9), designating any area of acoustic study, studies of sound environments in specific areas have been conducted using a combination of acoustic measurements taken in the area and qualitative surveys to obtain information about the perceptions of inhabitants in the sound environment (Porteous and Mastin 1985; Truax 2001). Feld (1982, 44–60) lived with the Kaluli people of the Papua New Guinea central plateau to study their song, poetry, weeping and dance performance. He found that bird song and the Kaluli understanding of birds as simultaneously creatures in the environment and spirit reflections of dead men and women greatly influenced their customs and beliefs. Feld found that Kaluli cosmology informed their interpretation of bird taxonomy. Certain bird species were grouped into families in the taxonomy not because of similarity in physical traits but because they played related roles in the Kaluli stories of the creation of the world. The role of birds in the Kaluli cosmology influenced their interpretation of bird song and behaviour which, in turn, was reflected in their cultural performance.

Sui (2000) argued that at least some geographical research in recent decades has been based on aural, rather than visual, metaphors and he proposed this trend to be an echo of several changes within geographical research, partly in response to extra-disciplinary factors: philosophical approaches critical of visual metaphors; the re-emergence or reinvigoration of aural/oral traditions because of the development of electronic communication media (McLuhan 1962); and the increased participation by women within geographical research in general and the emergence of distinctly feminist geographies in particular.

According to Sui (2000, 329), aural culture immerses its members in a mesh of acoustic events and knowledge interchange itself assumes the character of time-bound events rather than the character, assumed by knowledge in visual culture, of an object. According to Dyson (1995, 29), "*aurality* refers to the phenomenal and discursive terrain of sound. [. . .] The phenomenal invisibility, intangibility, multiplicity, and existential flux of sound challenges an understanding of the real based upon the visible, material, and enduring object." As Sui argued, the understanding of culture, society, and knowledge as currents of events implied by a re-emergence of aurality would suggest an alternative set of metaphors that could prove useful to geographical researchers. Such metaphors are evident, for example, in calls to recognize uncertainty in identities, those of researchers and subjects alike, as being constituted by the performance of those identities (Pratt 2000; Rose 1997)[1] and in calls to recognize the role of politics and power in the constitution, but not determination, of political identities and agency (Butler 1992). In these conceptions, the ongoing definition of identities, elsewhere framed as stable, seems more amenable to construction as a sequence of events rather than as reified facts. Moreover, the engagement implied by the ongoing performance of identities, between researcher and subject or among people who may fully or partially share a political identity, seems better characterized by aural (e.g., dialogue, conversation) than visual metaphors.

According to Sui, aural metaphors should be understood as co-existing alternatives to the dominant visual metaphors in use by Western culture generally and within geographical research in particular. Visual and aural metaphors all provide insights with inherent strengths and weaknesses and geographers should "develop a *both/and* stance" (Sui 2000, 336) toward them, deciding what to use and when based on an understanding of the inherent assumptions underlying them. Such a model of research would then balance the distanced, synoptic model of visually-acquired knowledge and the relational, "turning toward the Other" (ibid., 325) implied by listening.

Although aural space is, as described above, a mesh of events, visual space, despite the usual characterization as a distanced space of objects, is not static. Un-

[1] Although the performances Pratt discusses can be described as aural, Pratt describes her own process of understanding and incorporating this material into her research as a process of writing. As such, her work probably relies on a set of metaphors combining at least aural and textual elements and thus may highlight that Sui's visual/aural dichotomy is but one abstraction of the possibilities.

derstanding the visual perception of static objects, and the visual grouping and organizational principles of objects within a space has been one, but not the sole, focus of research (Hershenson 1998). Participation in a visual situation may involve either a moving observer or moving objects. A kinetic understanding of visual situations implies that vision must also be concerned with events and processes. The omission of these from the usual characterization of visual metaphors may then be an overstatement on the part of detractors of certain visualities[2] or it may be a learned perceptual emphasis from certain visualities. But, as will be discussed in Chapter 5 in more detail, a concern with process and events has been at least part of the impetus for the development of (visual) animated and interactive cartography. Thus, the temporal dimension of audio and the relational, event-oriented metaphors of aurality may align well with the focus on the visualization of processes demonstrated within dynamic cartography.

1.3 Predominance of Vision in Cartography

Maps and map-making in Western traditions are usually considered in terms that are consistent with their genealogy as visual, paper-based practices. But the technological basis of map-making has changed with the introduction of computer technology into the making of maps. Similarly, home computers and mobile technologies such as Global Positioning System (GPS) receivers and cell phones have been widely adopted for, among other purposes, the reception and use of maps distributed digitally. Distribution using computers connected to telecommunication networks such as the Internet and World Wide Web has been the most prominent of these digital methods (Peterson 2007) but the increasing adoption and

[2]For example, the characterization, critiqued by Jay (1988, 8), of all visualities as "a reifying male look that turned its targets into stone."

use of mobile phones and the multifunction designs being introduced in 'smart' phones are very likely eroding that predominance, at least for mainly locational uses of maps (as opposed to more data analytic uses). Augmenting the one-to-many distribution model of paper-based publishing, the many-to-many connectivity of these networks is also being used to enable the creation of maps and geographic data by loosely-coupled organizations, collaborating groups, and individuals all having common interests in pooled data (Goodchild 2007; Sui 2008).

Although the initial uses of computer technology in map-making were motivated mainly by desires to automate data collection and manipulation activities required in the production of paper maps, these transformations of the technological basis of map making and distribution have led to changes in the design and use of maps. More recent research examining the possibilities for using computer technology as a means of using and interpreting maps has resulted in the development of dynamic map products that allow a map user to modify the map display through interactions with controls designed into the interface of map-based artefacts. By endeavouring to thus transform maps, the map making community and the users of its artefacts can be seen as part of a larger contemporary process of media and technological convergence which has resulted in displacements and diminutions of existing analogue technologies in favour of digital formats, processes, and systems (e.g., film by digital video, audio tape by digital sound, paper based rendering and storage of illustration by digital graphics). The mapping community can be seen to be both actively defining the processes of tool and artefact redesign by innovating solutions to the high costs involved in collecting, processing, and storing geographic information, and responding to economic and technological trends that would have made a lack of process redefinition, at the least, very difficult as existing technologies became less mainstream.

Experiments incorporating sound with maps are certainly not new but the literature on the cartographic uses of sound is sparse and there is as yet little guidance concerning ways in which sound can be effectively used. Guidance concerning which purposes sound can most effectively serve has also been rare (cf. Krygier 1994). Mapping products and cartographic processes incorporating sound along with visual components (e.g., map graphics, images, and text), although being developed on an experimental basis by artists, multimedia designers and cartographers, still constitute a small proportion of contemporary mapping efforts. Recent Internet applications and World Wide Web page design technologies, such as Google Maps (Taylor 2005a), enable the use of audio as one among many forms of media annotations on maps. The resulting audio 'mash-up' maps are widely available through the Internet.

Examples of maps incorporating sound in conjunction with visual maps have been developed over at least the last two decades as an outgrowth of research into *animated* and *interactive* cartography and GIS. In these examples, sound has been used to provide narration explaining the function of a map or a multimedia interface (Harrower 2003; Krygier 1994; Monmonier 1992); to draw a user's attention to a visual map component (Harrower 2003); as an abstract language providing a set of variables (pitch, tempo, timbre, etc.) adding information to a thematic map without over-complicating the visual display (Fisher 1994; Krygier 1994); and as media to be accessed through the map, alone or as part of a video or an animation (Hu 2003).

Krygier (1994) examined possible uses of sound in conjunction with cartography as enabled by computer technology and, although he discussed the use of narration and mimetic sounds, he focused in particular on the definition of a set of abstract sound variables such as pitch, timbre, location, and duration that can be manipulated to encode data associated with visually-mapped geographic locations. Fisher (1994) demonstrated the use of one of these abstract sound variables (duration) to highlight data uncertainty within a cartographic display. Sound has been used as an added dimension in GIS to provide qualitative and quantitative information for studying noise abatement policies (Müller and Scharlach 2001; Servigne et al. 1999).

Multimedia cartography (Cartwright et al. 2007) incorporates audio and video clips, still images, animation, and text into dynamic cartographic products, often using the map as an interface providing access to the multimedia content. Examples include *Wula Na Lnuwe'katiyek*, an atlas of Mi'kmaq places and stories (Francis 1996); *Introduction to Iqaluit*, a set of maps combining imagery and place name pronunciations (Mouafo and Müller 2002); and *Gwich'in Place Name Map*, a set of maps organizing access to written stories, videos, and place name pronunciations concerning Gwich'in use of their territory (Gwich'in Social and Cultural Institute 2003).

Gartner (2004) discussed the use of handheld mobile Internet devices to assist pedestrian navigation by issuing redundant spoken-word directional commands that compensate for situations in which a user may be focused on other tasks or when circumstances make reading the screen difficult. Théberge (2005) argued for multimedia atlas designers to consider sound early in the overall process of atlas design because visual and acoustic design elements each affect the interpretation of the other. Levine et al. (2004) created a web application as part of *Baghdad San Francisco*, an installation incorporating map transpositions and geocaches transposed from Baghdad to San Francisco documenting the bombing of Iraq by US armed forces during the initial invasion of Baghdad in March 2003, in which an ambient audio recording derived from the bombing underscores the political message of the piece. Levine (2006) created an installation documenting the impact of the 1906 San Francisco Bay area earthquake on the city of Santa Rosa in which audio derived from seismic recordings of the earthquake accompanied visual distortions of a projected map image of the city. Thirion (2007) created a map-based data visualization of Madrid automobile traffic projecting an explicit authorial perspective onto the underlying traffic flow data through the use of a *traffic as noise* metaphor. Caquard et al. (2008) used sound within a cybercartographic atlas project to enhance and highlight the narrative character of an atlas and the potential of that narrative to encompass multiple perspectives.

These examples of cartographic sound design indicate alternative theoretical perspectives and approaches with, for example, Krygier (1994) adapting cognitive-cartographic perspectives on dynamic visual displays (DiBiase et al. 1992) to abstract sound use and Caquard et al. (2008) examining the theory of sound design for cinema to derive approaches for cartographic sound use.

Although these examples show that interest exists in the use of sound as part of map-based artefacts, this list of examples is tiny in comparison to the volume of silent maps available in books or on the World Wide Web. Furthermore, a review of maps available on the World Wide Web that incorporate sound, conducted during June 2010, shows that most include sound in such a way that it is possible to argue that the designers still think of the map itself as being a silent visual organizer for the sounds but not an audiovisual device on its own. See Appendix A for details. Tables A.1 and A.2 present results from the survey.

Table A.1 presents web-mapping applications found that incorporate sound(s) as part of an audiovisual map interface. Sounds, possibly multiple simultaneously, play while a user interacts with the visual map interface. There is no intermediate visual object, such as a pop-up display containing an audio player, that a user must manipulate in order to begin audio playback. The result is that the map itself appears as an audiovisual device, the source of visual and auditory signification.

By contrast, the web-mapping applications listed in Table A.2 use a map to organize sounds geographically but separate playback of those sounds from the visual map through the use of an intermediate visual device, such as a pop-up speech bubble overlaid on the map or a separate HTML page, providing playback controls and information about the audio. The result is the creation of a perceptual distance between the visual and auditory signification of the map application: the sounds are organized *by* the map but are not part *of* the map. This despite the description of these maps, varying somewhat across the projects, using terms such as 'sound map' or 'acoustic map.' More than 80% of the maps examined during this review (33 of 40) are designed using this *location of media* design approach. Of those using such a design, the majority offer access to recorded audio and other media such as images or videos through the use of a pop-up bubble that appears on the map when a map feature (usually a point) is clicked (21 of 33), occluding visibility of at least a portion of the map. The next most common design amongst these samples (10 of 33) provides access to the recorded audio and other media through a separate HTML page reached by clicking on mapped entities. Despite being described as a sound map, the map itself has disappeared before a user has access to any audio.

The majority of maps organizing while separating sounds from the map were created using common web map application programmer interfaces (API) such as the Google Maps API or the Bing Maps (formerly Microsoft Virtual Earth) API. Although examples and tutorials for adding media to maps using these APIs tend overwhelmingly to present media in pop-up bubbles or outside of the map, the API designs do not necessitate this. The APIs support the ability to run application-specific javascript functions when map features are clicked or when the cursor hovers over features. Thus direct triggering of audio playback as a side-effect of user interactions would be straightforward but seems to be rarely considered in designs using these APIs.

The remainder of the map applications listed in Table A.2 (2 of 33: *Cinco Cidades* and *London Sound Survey*), although they do not occlude or remove the map before audio playback starts, do not provide visual feedback that enables a user to understand the sounds heard as being associated with places on the map. *Cinco Cidades* is one of several applications reviewed that support simultaneous playback of multiple audio recordings and is one of three applications produced by *The Folk Songs Project*. Both *Folk Songs for the Five Points* and *Manchester: Peripheral*, by the same group, do provide a clear indication of the locations associated with audio recordings while they play, unlike *Cinco Cidades* which allows a user to select audio from the map but the visual map then retains no trace of where those sounds were found once they start to play. *London Sound Survey* supports the selection of a group of audio clips based on location classified according to a regular grid overlaid on London. Once a grid cell has been selected, all of the audio recordings associated with that location are loaded into a side panel and may be played alone or concurrently with other recordings for that cell but the map retains no indication of the grid cell with which these recordings are associated.

This classification is not intended as criticism of these designs. The only important criteria for these applications is whether or not they support the objectives derived from their projects' goals. For example, the spatial organization of sounds for acoustic ecology projects (e.g., *Toronto Acoustic Ecology*) or for collaboration and sharing of audio amongst independent film producers (e.g., *The Smalls Street Sounds*) may meet all of the requirements of these projects and the availability of web mapping APIs that provide a simple means of meeting those requirements probably explains the common design approaches taken in many of these projects.

For this thesis, the main conclusion drawn from this review and the main purpose of the review was to highlight that design possibilities for utilizing sound and for creating complementary visual and auditory designs for web-mapping applications are not yet nearly exhausted. There is still much scope for research on these and related topics. As argued above, the design possibilities of even popular web-mapping APIs are rarely explored. Although the *location of media* design pattern may be adequate and easy to use for certain projects, I do find it interesting how remote sound is from at least some of these maps and how strong the visual bias remains in the designs, considering that these are generally described by labels such as "sound map." Beyond issues of technology, this bias reflects entrenched patterns of thought concerning maps and sound.

The argument thus far has been that sound is usually not considered in the design of maps but there is also the possibility that it is being considered more often than I am allowing but it is being rejected, possibly due to preconceived notions concerning the value of sound as part of interactive computer applications, in general, and interactive maps in particular.[3] Acknowledging this possibility does not discount the argument that the use of sound as part of audiovisual cartography warrants at least more research than it currently gets. Starting from a discussion of the reasons and motivations for using sound and building on that to develop a theoretical framework for the use of sound in audiovisual cartography, this thesis attempts to develop some strategies and approaches for designing maps to include

[3]Perhaps unsurprisingly, it is not difficult to find dislike expressed for the use of sound in maps. As an example, Woodruff (2009) enumerates "annoying sound effects" as a possible problem with animated maps, along with "dizzying excessive motion, poorly or over-designed interfaces [...] and perhaps a few dancing bananas."

sound. This requires the creation of models for the distribution and use of sound as part of audiovisual maps, the development of tools and software implementations that meet the requirements of such models, the development and use of map examples using those tools, and the assessment of user responses to the maps created. Only through combined theoretical and applied research into the possibilities for designing audiovisual maps can the potential value and applicability of sound for such use be properly assessed.

2

Changing Media Trends and Use of Maps

This chapter discusses the media environment in which I am arguing that sound should more frequently be considered as a component of maps. Two questions deserve some discussion. First, what do we know about the accessibility of information and communication technologies (ICT), the degree to which these technologies have diffused throughout national and global populations, through which people would be able make use of such maps and what do we know about the existence of knowledge, abilities, and desires to actually use ICT within populations that have access to them? Second, what do we know about how ICT are used and consequently about how the use of audiovisual maps could be accommodated in conjunction with existing uses of these technologies. No unequivocal answers can be given for either of these questions although governments, industry groups, and researchers do attempt to estimate the diffusion of ICT both within nations and globally. Less attention has been focused on how the technologies are used than on the physical diffusion of the technologies because understanding the phenomena of use is inherently more difficult (Papper et al. 2004). To briefly examine the availability of ICT and trends in uses of these technologies to access the Internet for different purposes, I will present data derived from national surveys in

Canada, followed by brief comparisons of global regions.

Diffusion of ICT across regions and countries is known to vary according to relative wealth and issues of policy and governance (Castells 2001; Drori 2005, 2010). In addition, opportunities to use ICT, even within regions in which the technologies are relatively plentiful, along with levels of ability and interest in the use of ICT are differentiated across socio-economic and political strata. Rather than a single *digital divide* separating global scale geographic regions by levels of physical diffusion of ICT, there are a set of multidimensional gaps that differentiate access to and usage of ICT at all scales. Variations in usage and benefits deriving from use can be correlated to differences of age, gender, education, wealth, ethnicity, location, and occupation (Castells 2001; Drori 2005, 2010; van Dijk and Hacker 2003; van Dijk 2005; Warf 2001). As van Dijk (2005, 2) argued, gaps in usage of and derivation of benefits from the use of ICT have persisted, even in places where issues of physical access to the technologies have been widely addressed, and have sometimes exacerbated existing disparities: "where most people are motivated to gain access and physical access is spreading, differences in skill and usage come forward."

As shown in Tables 2.1 and 2.2[1], surveys of Internet use amongst Canadians by Statistics Canada show that general use of the Internet increased between 1997 and 2009 and that over that decade a greater percentage of the population was using the Internet to access information and entertainment. The information and entertainment accessed was being obtained in a range of formats: written texts, audio-only formats such as online radio streams, and video formats that often include

[1]The methodology for Statistics Canada's survey of Internet use changed for the 2005 results included in Table 2.2. The data shown in Tables 2.1 and 2.2 cannot be directly compared. Furthermore, the target populations surveyed for the data shown in Table 2.2 changed during the period for which data has been collected with people aged 18 and over surveyed in 2005 and people aged 16 and over surveyed for later years.

imagery, text, and sound. Although these data show a general growth in use of the Internet and a broadening of purposes for which people were using the Internet, the survey questions asked to generate these data do not reflect usage that could be considered regular or consistent, much less usage that could be considered frequent. Table 2.1 shows Internet usage reported by households between 1997 and 2003 for all households, with the criteria that one member used the Internet once during the preceding year. Table 2.2 shows the percentage of Canadians using the Internet, by selected activity for the years 2005, 2007, and 2009. Again although the data show that an increasing percentage of Canadians accessed the Internet over time, for a range of purposes during the period of these surveys, the frequency of use for an individual to be included as a user of the Internet is very low: once during the year. However, according to a summary of results from Statistics Canada's 2009 survey of Internet use (Statistics Canada 2010) the frequency and duration of use by regular Internet users are also increasing. This summary argued that 96% of Canadian Internet users accessed the Internet from home and, of these home users:

> 75% went online every day during a typical month, up from 68% in 2007. In 2009, 55% were online for five hours or more during a typical week, up from 49% in 2007 (ibid., 3).

Based on a 2007 survey conducted in Canada, Zamaria and Fletcher (2008) argued that differences in the likelihood of an individual using the Internet were related to a number of socio-economic factors including income, education, professional status, age, mother tongue, gender, and location. Levels of income, education, and professional status were positively correlated with increasing likelihood of Internet use and access although the gaps in access and frequency of use had narrowed between 2004 and 2007 (ibid., 6). Age was inversely correlated with

Table 2.1: Percentage of Canadian households reporting Internet usage, by selected activities and year: 1997–2003. For activities other than "Internet access," only activities conducted from home are included (Statistics Canada 2005). A household is said to use the Internet if at least one member of the household accessed the Internet for the specified purpose at least once during the year preceding the survey. "Regular Use" households are those reporting that at least one member accessed the Internet for some purpose, from any location, at least once a month (Statistics Canada 2006). Source: Statistics Canada, CANSIM tables 358-0006 and 358-0017†).

	Selected Activities	1997	1998	1999	2000	2001	2002	2003
Percentage of all households	E-mail	13.3	19.3	26.3	37.4	46.1	48.9	52.1
	General browsing	13.6	17.6	24.3	36.2	44.3	46.1	48.5
	Playing games		7.8	12.3	18.2	24.4	25.7	27.9
	Obtain and save music			7.8	17.8	23.3	24.3	20.6
	Listen to the radio			5.0	9.3	12.3	12.3	13.1
	Find sports related information				17.3	22.1	23.8	24.6
	View the news				20.4	26.2	27.2	30.2
Regular use households, percentage of all	Internet access, any location†	29.4	35.9	41.8	51.3	60.2	61.6	64.2

access to and use of the Internet with 96% of Canadians aged 12–17 years using the Internet compared to about half of Canadians aged 60 and over (ibid., 42). Current use of the Internet (during the 3 months prior to the survey) by Anglophones (82%) was greater than that by Francophones (67%) but this gap was reduced to 8% (Anglophone use was 85% compared to Francophone use at 77%) if use of the Internet during the previous year was compared (ibid., 42). The gender gap in access to and use of the Internet had narrowed in Canada to within the margin of statistical error with 80% of all males active online compared to 76% of all females (ibid., 43). But males were spending more time online (ibid., 59) and were engaging in a wider diversity of online activities (ibid., 162). In addition, among Canadians 60 years and older a greater percentage of males (56%) than females (46%) were using the Internet (ibid., 43). Communication technologies providing access to the

Table 2.2: Percent of all Canadian individuals using the Internet at least once within the last year, by selected activities: 2005–2009. For activities other than "Internet access," only activities conducted from home are included (Statistics Canada 2009a). The survey includes individuals 18 years and older for 2005 and 16 years and older for 2007 and later. Source: Statistics Canada, CANSIM tables 358-0130 and 358-0122† (Statistics Canada 2009a, b).

Selected Activities	2005	2007	2009
Internet access, any location†	67.9	73.2	80.3
E-mail	55.6	63.1	71.7
General browsing (surfing)	51.2	52.1	59.9
Obtaining weather reports or road conditions	40.5	47.9	57.5
Viewing the news or sports	37.6	43.7	52.2
Education, training or school work	26.1	34.0	38.8
Obtaining or saving music	22.3	30.5	35.9
Playing games	23.5	26.5	32.4
Listening to the radio over the Internet	15.9	19.3	24.5
Downloading or watching television	5.2	10.8	19.1
Downloading or watching a movie	5.0	8.6	15.3

Internet, and especially high-bandwidth access, were still less available in smaller communities and regional differences in the percent of the population using the Internet were still evident (ibid., 70).

The Telecommunication Development Bureau of the International Telecommunication Union (ITU-BDT 2009a, b, c, d, e, f) argued that all global regions have experienced growth in the number of Internet users as a percentage of the population during the years 2002–2008,[2] although conspicuous national use rate gaps remain. For example, the ITU-BDT estimated that 44% of people in the Americas used the Internet but noted that "[p]enetration in the United States and Canada, with two out of every three people on line [sic], is almost two and half times higher than the Latin America and the Caribbean average" (ITU-BDT 2009b, 17). National use rate gaps exist across all of the regions for which the ITU-BDT reports with, for

[2] For all regions except Asia and the Pacific (AP), the ITU-BDT estimates discussed in these reports are for end of year 2003 and 2008 (ITU-BDT 2009a, b, c, e, f). For AP, the estimates are for end of year 2002 and 2007 (ITU-BDT 2009d).

example, use rate estimates for the most recent years reported,[2] as percentages of the population, ranging between 0.2% (Sierre Leone) and 37.8% (Seychelles) in African countries (ITU-BDT 2009a, 60), between 0.1% (Myanmar) and 70% (New Zealand) across Asia and the Pacific (ITU-BDT 2009d, 50), and between 23.9% (Albania) and 90.6% (Iceland) for Europe (ITU-BDT 2009f, 70). These data do not address frequency of use within these disparate populations and these would be expected to vary as dramatically as the access rates.

The ITU-BDT further argued that regions with lower levels of use are experiencing higher growth rates. Although the higher growth rates experienced by some regions indicate that, in terms of the number of Internet users as a percentage of population, the regional gaps may be beginning to narrow, several factors indicate that a straightforward closing of the gap is unlikely or will still require a considerable time to occur. First, the gaps between those areas with the greatest levels of use (especially North America and Western Europe) and those experiencing the highest rates of usage growth are still very wide. Second, technology adoption within a population, measured as a percentage over time, is often non-linear with higher rates of growth experienced in early phases of adoption and slower rates later. This non-linearity suggests that regions experiencing high growth rates now are likely to experience slower growth in the future resulting in regional gaps remaining yet for significant periods (Robison and Crenshaw 2010). Third, certain regions (e.g., Africa and the Commonwealth of Independent States) experienced even greater growth in mobile cellular subscriptions during the years reported,[2] partly because this provides communication services without requiring the creation of wired infrastructure (ITU-BDT 2009a, b, c, d, e, f). Therefore, if the national gaps in Internet usage discussed by the ITU-BDT do close, the technological means, or at least the mix of technologies, by which the Internet is used in some of these high-growth regions may not be identical to that of the current high-use regions.

The ITU-BDT reports (ITU-BDT 2009a, b, c, d, e, f) only minimally addressed issues of unequal access to ICT within the regions or countries covered. For example, gender gaps were only addressed in the report on the Americas where it was argued that there remains only a "minimal gender gap" (ITU-BDT 2009b, 30), although among the countries for which that report included gender comparisons it showed gaps of as much as 8%.[3]

With regard to the purposes for which people use ICT and the Internet, other surveys and market reports for Canada provide additional details, focusing on the use of a variety of media types including texts, music, video, and games.

Text: A survey conducted in Canada on behalf of the federal Ministry of Canadian Heritage during 2005 found that online reading was becoming an increasingly common activity (Créatec + 2005). The study reported that 25% of respondents had subscribed in the previous year to at least one newspaper, newsletter, or magazine that allowed publications to be read online. Zamaria and Fletcher (2008) reported that 46% of Canadian Internet users said they had read an online newspaper or magazine (24% did so at least weekly) and 15% said they had read a book online (only 4% did so at least weekly).

Music: Walsh (2010) reported that in Canada digital music sales are increasing over time, although still only representing 14% of total album sales (by units sold) in 2009, and sales of physical albums such as compact discs continued in 2009 a declining trend that began almost a decade ago. The International Federation of Phonographic Industries (IFPI 2010), an international recording sector association, reported that through 2009 digital music sales world wide, including download

[3]In Peru, 33% percent of males used the Internet compared to 25% of females

sales and online streaming subscriptions, continued to grow and now represent more than a quarter of music revenues world wide and 40% of US music revenues.[4] Since 2008, Apple's online iTunes store which sells music only in downloadable digital formats has been the leading seller of albums in the USA, when comparing all product sales including those of physical formats (Quinn and Chmielewski 2008).

Video: Online video sites are becoming increasingly popular in Canada. Zamaria and Fletcher (2008) argued that 40% of Canadian Internet users reported that they at least occasionally[5] used the Internet to download or watch videos and among 12–17 year olds the percentage participating in this activity rose to 79%.

Gaming: Zamaria and Fletcher (2008) argued that 43% of all Canadian adult Internet users in 2007 played online games with 25% reporting that they played at least weekly and among 12–17 year old Canadian Internet users the participation rate increased to 79% with 39% reporting that they played at least weekly.

Comparing the results of the studies cited above is difficult because of differences in methods, questions asked, and when surveys were conducted. Thus it may be difficult to draw conclusions from these studies with a high degree of precision but some general observations can be made:

- Diffusion of ICT is occurring in all regions, the technologies appear to be becoming increasingly accessible but are by no means ubiquitous.

[4]IFPI (2010) provides an overview of its members' sales while arguing for legislative controls to prevent non-sanctioned uses of recorded music. It is difficult to estimate overall activity and growth levels for non-sanctioned uses of recorded material: industry organizations like IFPI estimate very high levels of non-sanctioned uses and many Internet Service Providers through whose equipment the non-sanctioned uses could be detected are uninterested in being seen to monitor their customers' online activities, although legislation in some countries (e.g., France) has begun to impel them to do so.

[5]The response indicating the least frequent possible use of the Internet to download or watch videos yet indicating some level of participation was "less than monthly."

- Beyond physical access, there are a variety of socio-economic factors that differentiate the opportunities for people to productively use and benefit from ICT, both within relatively *high-use* countries like Canada and the USA[6] and even more markedly when comparing across global regions.
- Among some groups, and especially within regions in which ICT are widely available, very high rates of ICT use exist.

In addition, although not yet discussed above, *media multitasking*, simultaneous use of multiple media such as listening to music while reading, is common among high-use groups. As Roberts and Foehr (2008, 11) argued, "American youth are awash in media. [...] They spend more time with media than any single activity other than sleeping, with the average American eight- to eighteen-year-old reporting more than six hours of daily media use." This discussion next examines phenomena of media multitasking in more detail because this almost certainly has implications for how some people would use maps distributed over the Internet but has received little attention.

2.1 Media Multitasking

The phenomenon of media multitasking deserves attention. Although the observations in the preceding section, drawn from studies of ICT and Internet usage, mainly discuss a single type of media at a time, recent studies have noted or explicitly examined the simultaneous use of multiple media. Papper et al. (2004,

[6]For brevity, results from Internet usage surveys conducted within the United States have not been discussed. Although details differ, broadly similar results to those discussed for Canada have been identified during the same periods covered by the Canadian surveys, including general growth of ICT diffusion and use and the existence of multidimensional gaps in rates of usage across geographic, social, cultural, educational, and economic differences (Horrigan 2006; Lenhart et al. 2008; Madden 2009; Papper et al. 2004; Rainie 2010).

39) studied media use within the population of 'Middletown', USA and argued, based in part on discrepencies between self-reported and observed times participants spent using and exposed to different media types, that

> [p]eople spend almost a quarter of their media day with two or more media, and much of that multiple use appears to go unnoticed by the people who do it. It is clear some people engage in more multitasking than others, and certain media are more prone toward being part of multitasking than others.

The suggestion that, at least among some groups, concurrent attention to multiple media is common or near rampant is remarkable for the lack of attention this has received in studies of online or digital map use since, as argued by Peterson (2007), the Internet became the predominant distribution channel for maps. Although beyond the scope of this thesis, the possibilities that maps intended by their designers to be used on their own may be part of a concurrent mix of attention-seeking media has direct implications for the focus of this research as well as for that related to all digitally distributed and used maps. As will be discussed below, and as suggested in the above quote by Papper et al., audio is often used as a secondary medium while a user is primarily focused on other media. It is unknown whether or not introducing sound as part of an audiovisual cartography will create additional difficulties or create the potential to alleviate a problem that is largely unknown among digital map designers. If a map's sound design competes for the attention of a user intending to listen to something else then this could be problematic. By contrast, if you accept that what people are listening to can influence their interpretation of visual media, a topic I discuss in Chapter 4 with respect to the functions of sound in film and games, then I may also be advocating an approach to maps that allows those making maps to make assumptions about what *is* being listened to while the map is in use.

Based on surveys and observations of participants, Papper et al. (2004) argued that by far the most common forms of media multitasking involved use of television along with other media. But based on observation methods whereby researchers designated the media that was predominantly engaging participant attention as the *primary* media, they argued that multitasking, especially with television and use of a computer, tends to be asymmetric. For example, when participants were watching television as their primary activity they were usually doing that without other media but the television was often on as a background to other activities such as talking on the phone or reading. According to the observation studies, relatively few participants watched TV while using the computer but those that did would do so for significant periods of time, averaging more than an hour per day (ibid., 29).

Similarly, use of the computer while multitasking with media other than television was found to be asymmetric. Participants commonly used secondary media while their primary focus was what they were doing with a computer. Papper et al. (2004, 29) argued that "over two-thirds (67.1 percent) of computer users also watch video or listen to audio at least some of the time; 43.4 percent of the time that the computer is in use, respondents also used a video or audio medium." It was less common for participants to use the computer as a secondary activity while attending to other media.

Roberts and Foehr (2008, 19), examining a broad range of media use among children and youth in the USA, argued that media have become so pervasive in the lives of young people that studies should distinguish between *media use*, "the amount of time young people devote to all media" (i.e., person hours), and *media exposure*, a measure of the "media content encountered by young people expressed in units of time" (i.e., medium hours for each type). They found that concurrent media use had become such a common practice among youth, with 81% reporting that at least some of their media use involved concurrent exposure to more than one media (ibid., 28), that enumeration of total media exposure across types of media to which a person had access would often significantly exceed the person's media use. They argued that, although total media exposure had increased between 1999 and 2004, the increase was overwhelmingly the result of increased concurrent use of media: "total media exposure increased, media multitasking increased, total use remained relatively constant, and there is little evidence that any medium — but especially television — is being replaced" (ibid., 19–20).[7] Calling computers "media multitasking stations", Roberts and Foehr found that most of the time youths spent working with a computer also involved secondary activities, most often involving additional media, and most often involving another form of computer media (ibid., 30). They speculated that certain types of media lend themselves more readily to multitasking and that certain combinations (e.g., reading while listening to music) seem intuitively more compatible than others (e.g., watching TV with the sound on while listening to the radio).

Zamaria and Fletcher (2008, 206–207), based on phone surveys, reported that 76% of Canadian Internet users would engage in more than one activity while using the Internet and 36% would do so often. Furthermore, they argued that the likelihood of engaging in concurrent activities while using the Internet was highest among those groups that spend the greatest average time online, especially young adults, students, and youth. The most common activities conducted simultaneously with use of the Internet included use of the telephone, listening to music, watching television, and listening to radio. "Across all Internet users, 18% of time

[7] A survey by Statistics Canada of social time use practices *has* argued that the proportion of the Canadian population spending leisure time watching television diminished slightly during 1998–2010 because of increased use of computers, mainly among younger users (Statistics Canada 2011).

on the computer is also spent watching television, while more than half the time online (54%) is also spent listening to music."

Although Zamaria and Fletcher (2008, 207) suggest that the phenomenon of concurrent media use is new or qualitatively different with ICT and the Internet, there is a good chance that the difference is as much in the way in which researchers are beginning to understand and measure media use and exposure. Earlier technologies and their social uses demonstrated acceptance of media multitasking when, for example, audience antipathy to silent exhibition of films spurred exhibitors to hire musicians or lecturers or use recorded music as film accompaniments (Altman 1992a; Berg 1976; Cavalcanti 1939), as will be discussed in more detail in Chapter 4. Transistor radios and Sony Walkman portable stereos and more recently Apple iPods, have enabled listening to be used as a strategy to assert emotional and social preferences within initially private but increasingly public spheres (Bull 2000, 2007; Tacchi 2002). Although much has been made of the effects of mobility on media use (Bull 2000, 2007), intrinsic to these devices has been their ubiquity and the multitasking implied by that presence. Writing about transistor radios, Tacchi (2002, 242) declared his starting point to be "the idea that radio sound creates a textured 'soundscape' in the home, within which people move around and live their daily lives." Chen (1998) argued that for some college student Walkman users their personal stereo was almost always present, accompanying them in social and private situations, sometimes establishing their situation as private, and often used to pace their work and study routines. Kassabian (2008, 120), by contrast, argued that "the particular problematic raised by simultaneous listening, and thereby multitasking, is that of attention." How do we understand media use that is continuously interrupted "by listening with part of an ear to music while you are chatting, shopping, reading, working, driving" (ibid., 121)?

The near continuous presence of the television, at least in some households, as a background to many other activities probably predates the presence of computers by decades. It seems clear that many people, especially children and youth, at least in Canada and the USA are learning to multitask in their use of media. But suggestions that youth show a greater tendency to multitask with ICT and the Internet may be simply an indication of their greater level of familiarity with that particular set of technologies. It seems clear that this is not the first generation to engage in media multitasking. Whether or not there are aspects of the use of ICT that make the current phenomena qualitatively different remains an open question.

The shift in the recognition of these phenomena is significant and probably deserves attention among cartographic researchers. As Olson (2004, 58) argued, in reference to the 'Middletown' studies, "it is clear from the study that users of media cobble together textual experiences that come from different media sources at the same time, a contradictory, nonlinear, and incoherent experience." What does this realization concerning the prevalence of media multitasking imply for the use of maps and the methods of cartographic research in general? This suggests a need to understand why, when, and how map users multitask when working with a map and the answers are going to be partial and complex. Part of the assessment of design requirements for a mapping project should, when possible, include an assessment of when and where map users will commonly use the map and what other media they are most likely to be using concurrently. Due to difficulties in identifying even the most common tasks to be conducted with a proposed map design, except in cases involving only the most specialized of map types, this is rarely done. For many general purpose maps and educational atlases, the answers are unlikely to be precise, resulting in the need for designs that are flexible in accommodating the possibilities of multitasking.

2.2 Implications for Audiovisual Cartography

Despite the World Wide Web now being the predominant means of distributing and using maps, access to and use of technologies enabling such use is very uneven at multiple scales. The prospect of digital map design, distribution, and use, let alone those same activities in relation to audiovisual maps, must then be considered along with the realization that, in certain regions and for certain groups in all regions, access to such products and processes remains constrained.

Considerations concerning simultaneous use of multiple media by prospective map users and what this means for theories of map design and use are potentially problematic. Cartographic researchers during the twentieth century, following Robinson (1952), have argued for investigations that bring together understandings of the cognitive processing of maps and understandings of data transformations and methods of representation to improve the learning potential and functional effectiveness of maps. Other theoretical perspectives have argued for a greater appreciation of the social contexts within which maps are produced, distributed, and used (Harley 1989; Jacob 2006; Wood and Fels 1986). Without rejecting any of the perspectives that would fall within the broad outline of either of these oversimplified characterizations, it is clear that each general approach is challenged by the current lack of understanding that exists concerning the contexts in which maps are used and the rapidly changing technological base upon which maps are now predominantly distributed and used. Studying map use with methodologies that isolate the map from other media activities, an approach which I must admit this research also adopted, is only valid if indeed map users normally exclude other media while focusing on a map. It is difficult to understand individual instances of media use based on averages and aggregated summaries of behaviour produced by studies of the type outlined in this chapter (general use and multitasking studies) but it may be reasonable to anticipate that map use within different contexts (e.g., use at home as compared to use in an office) may vary greatly.

Although wide variations in habits of media use and exposure undoubtedly exist, it seems clear that in regions in which ICT and the Internet are widely available, certain groups make very extensive use of ICT to access, use and produce media. In regions without widespread adoption of computers and Internet use it is reasonable to question whether or in what form such adoption will take place. In such regions, it may be reasonable to guess that cell phones or multifunction mobile devices may become the most widespread form of ICT or, indeed, it may be reasonable to guess that existing technologies such as television or radio continue to predominate. The importance of sound in the use of many mobile devices, if such devices do predominate in certain regions, could well make some of these a good focus for future extensions of this research, merging ideas developed here with those of telecartography (Gartner 2005).

3

Audiovisual Cartography as Digital Media

Jacob (2006, 100), in the process of defining maps, argued that a map is "a mediation between two mental images: that of its producer and that which its viewer will retain in the moment following its consultation." Maps provide form to spatial data, whether intended for private examination or public presentation (MacEachren 1994) and regardless of the relations between that data and any possible lived world of experience in terms of space, time, intent, or truthfulness. By giving form to data, a map shapes meaning and communicates, hopefully something about that data and, moreover, something useful about that data, but the map may also indicate something about the maker of the map and the purpose for which the map is being created. Continuing Jacob's (2006, 100) argument, "a map is never an isolated object independent of a desire to communicate, of the transmission of knowledge, and of a semiotic intent in the broad sense of the term." By focusing on the use of maps in electronic form distributed via information and communication technologies (ICT), I am considering maps as *digital media*.

In the above, I have been implicitly using the term *media* to refer to any type of communicative channel through which a person can access information. This is a common way to think about media and such a definition is entirely consistent with how the use of maps has been discussed. But another equally appropriate definition would be McLuhan's (1964) that considers media to be any extension of human capabilities. To McLuhan, media are those technologies and/or practices that extend human capabilities or that through their use by humans increase the pace or scale at which humans are able to operate in their lives. Communication technologies such as television, radio, and telephones all obviously fit this definition but, for McLuhan, so did the electric light by allowing people to operate in environments where they would otherwise be hampered by darkness. Maps, in their role as collectors and organizers of information, extend the spatial understanding of people, even for areas or spatial circumstances with which they have no direct knowledge, and thus fit McLuhan's definition as well. McLuhan (1964, 8) understood media to be complex, nested structures.

An alternative, and also applicable, definition of media that falls somewhere in between 'communication channel' and McLuhan's would be that of Gitelman (2006, 7) who defines media as "socially realized structures of communication, where structures include both technological forms and their associated protocols, and where communication is a cultural practice, a ritualized collocation of different people on the same mental map, sharing or engaged with popular ontologies of representation." This definition gives something of the sense of media as the outcome of intertwined cultural, economic, and technological practices where cultural practice includes accepted ways in which media are used in specific contexts. For Gitelman, the protocols associated with media include social behaviours developed around the uses of a medium such as the range of contextual greetings used at the beginning of a phone call. Podcasting, and the acceptance of podcasts as media by consumers, growing out of trends dating back at least to the popularity of earlier portable audio systems such as the Sony Walkman, can be seen as a cultural initiative driven by individualism, mobility, and a desire for control of personal listening habits (Bull 2000) more than as an inevitable technological progression encapsulating miniaturization, digital audio technology, file formats, and networked distribution. The technological developments and a base level of competence with the technology among potential users are necessary but not sufficient.

Sui and Goodchild (2001) argued that increasing availability of geographic information on the Internet and the growing predominance of communicative functions, among the possible roles that GIS could fulfill in its societal uses, meant that it would be valid to study GIS as media. They later applied McLuhan's (1988) tetradic analysis of media to outline an approach to understanding the societal effects of GIS, arguing that such an analysis would be useful in counteracting the prevalent technological determinism of much discussion of GIS and could be applied to guide future developments of and thinking about appropriate applications of GIS (Sui and Goodchild 2003).

For the purposes of this thesis, audiovisual cartography, which clearly overlaps with communicative aspects of GIS stressed by Sui and Goodchild, could be usefully understood as a complex of nested, digital media, some of which may have, to greater or lesser degrees and broadly understood, intrinsically spatial characteristics and others of which may be aspatial taken in isolation but nonetheless may be organized spatially within an audiovisual map or GIS. Ultimately it is up to the user of such a complex of materials to decide whether or not they understand it as a medium or as a multitude of media. The focus throughout this thesis will be on the use of audiovisual mapping and the embedding within it of digital audio, the information the audio conveys on its own, and the reworking of digital audio through spatial organization applied within an interactive visual map or an atlas comprised of multiple such maps.

In beginning to think about maps, and especially maps incorporating dynamic behaviours such as animation, interactive responses to gestures by users, and the playback of audio, it will be important to consider the characteristics of maps as malleable artefacts and, in the context of maps distributed and used electronically over distributed networks, this means thinking of them as programmable digital artefacts. Manovich (2001, 27) proposed that digital media share five "general tendencies", each of which is discussed briefly below: *numerical representation*, *modularity*, *automation*, *variability*, and *transcoding*.

Numerical Representation (ibid., 27–30): Products of digital media can be described functionally by a set of numerical values modelling elementary components of the media artefact as it will be reproduced when used or experienced. For example, a digital image may be represented as a grid of value sets defining colour and brightness (under a variety of colour models) for each cell in the grid. Numerical representation makes digital media artefacts amenable to numerical manipulation, allowing for example all of the grid cells in an image to be converted from colour to grey scale.

Modularity (ibid., 30–31): The repetitive functional definition of a media artefact, as outlined above for image grid cells, is common to all digital media with each artefact defined in terms of the quantitative specifications of sub-elements and known relationships between those elements. For example, a digital media image is made of many samples in which the relation between those samples is understood to be the spatial arrangement of the grid samples in two dimensions whereas a digital media audio recording consists of numerical representations of instantaneous audio samples in which the relation is understood to be the temporal order of the samples and the delays, usually constant, between taking samples.[1]

Automation (ibid., 32–36): Partly as a consequence of the two preceding characteristics, the production and use of digital media objects in the well-defined format specific to each can be automated. Indeed the ease of production of digital media and the resulting volumes of stored digital media has created the need for new methods of organizing, searching through, and retrieving digital media (ibid., 35).

Variability (ibid., 36–45): The separation of the production, storage, and representation of media objects across time and space, all in digital form, can allow for new representations to be developed subsequent to the creation of the original media object. This characteristic is well understood in digital map production with geographical databases and digital maps separating the dual roles of data storage and representation, previously both performed by paper maps. Digital representations, based on the stored digital data, may take different forms as required and could be in any of multiple possible graphic, audible, haptic, or other sensory forms and could represent the original data in isolation or in relation to other, possibly independently sourced, data sets.

Transcoding (ibid., 45–48): Manovich argued that through computerization of production, manipulation, and access, media objects, although retaining their properties as recognizable cultural artefacts, now also follow "the established conventions of the computer's organization of data" (ibid., 45). This requires, Manovich argued, an understanding of the digital media object as operating on two levels simultaneously: the *cultural* layer, in which recognizable media forms

[1]With the use of digital audio technologies, sound is 'recorded' by sampling, at a known frequency, the output of an analog-to-digital converter that is detecting and processing the source sound. The process has thus become known as 'sampling' audio and the resultant digital artefact, the stored sound, is referred to as 'sampled audio'. I will use the terms *recorded audio, recordings,* and *sampled audio* interchangeably to refer to objects of this type. *Samples* are single instantaneous measurements of the input analog signal, generally waveform amplitude for audio, many of which are required to create a useful fragment of sampled audio.

and genres such as popular music, newspapers, and movies are presented using the computer's interface, and the *computer* layer, in which the media objects are transmitted, stored, and modified by the actions of software and hardware as would any other distinct digital artefact. This implies the separation of data encoding the specifics of each digital artefact and the software, systems, and knowledge inscribing what they represent and how they can be manipulated and used.

Importantly this dual understanding of digital media artefacts includes circulations between layers not only in transformations of modular, numerical representations into recognizable cultural objects but also in conceptual transcodings from the computer to the cultural layer. This conceptual transcoding involves processes by which computer layer attributes and concepts, understood to be continuously varying and the result themselves of ongoing socio-economic processes, get adopted and extended within cultural processes. For example, podcasts are often promoted as convenient, on-demand entertainment alternatives to scheduled broadcasts. As such, 'podcast' has been appropriated within cultural contexts as shorthand for the cultural good while masking the actuality of a relatively complex process entailing economic exchange, intellectual property rights, and assumptions that intended listeners possess a base level of digital savvy sufficient to comprehend the production and depositing of a digital audio recording by an individual or organization for subsequent copying via a web application to a mobile device, possibly through one or more intermediary digital systems.

Manovich's concept of transcoding is not unique in identifying separation between data and software, and malleability in the resultant experiences, as significant characteristics of digital media. Murray (1997) referred to similar concepts as 'procedurality' while Bolter (2007) discussed 'code' and 'interface' as the underlying algorithmic and surface representational layers, respectively.

Manovich (2001, 96-100) further argued that when considering digital display technologies, such as circular-sweep radar screens or sequential scan televisions, there is a need to reconceptualize images in term of the underlying processes of production, transmission, and reproduction. For Manovich, time plays a fundamentally different role in both the production and representation of the captured image, as compared to a static image such as a photograph, with different portions of the screen showing image fragments from different instances in time. These digital displays create the illusion of a static image through the differences in the rate of change of observed objects and the update rates of the image production and reproduction procedures.[2]

Animated, and possibly interactive, cartographies and certainly audiovisual cartographies require that a map be understood as a processual interface. Animated maps most commonly represent spatial phenomena occurring over some period of time by scaling a representational time frame to create a dynamic display but can also represent alternative spatial criteria for a region by juxtaposing them through time. As such, map display updates must be understood in terms of the transformation of phenomenal time as it relates to the dynamically changing images.[3]

Sound is intrinsically a process-oriented phenomenon. Without events there is no sound. When an event creates sound, it plays out over time. Therefore, introducing it into dynamic cartography brings with it the necessity of managing sound cues as timed events in conjunction with considerations such as the mean-

[2]By contrast, Hansen (2004, 10-11) argued, following Bergson (1988), that embodied perception always precedes the creation of an apparently originating enframed image. "No matter how 'black-boxed' an image technology (or technical frame) may seem, there will always have been embodied perception at/as its origin." Although taking a very different conceptual approach to what the process of interpretation of images by a person would be, Hansen also stressed the importance of understanding digital media as processual (ibid., 9).

[3]Animated maps are discussed more fully in Chapter 5.

ings, connotations, and tone conveyed through the audio selections. When do they start and what duration are they? How do changes in the sound relate the visual data? If part of an interactive display, how do acoustic variations relate to the gestures of a user? In the remainder of this chapter, I discuss sound in audiovisual cartography using these perspectives on digital media.

3.1 Understanding Audiovisual Maps as Processes

Although map distribution over the Internet initially concentrated on the delivery of static map images and standardized specifications for mapping services on the Internet such as the Open Geospatial Consortium (OGC) Web Map Service (WMS) specification (OGC 2002) were oriented to that objective, more recent efforts and specifications have emphasized the programmability of the underlying data (e.g., OGC 2005). Map applications are emerging as hybrid media forms, for example combining spatial organization and real-time forum-based communication through crowd-sourced information contributions (Miller 2006).

Studying audiovisual maps for use through the Internet as digital media, the procedural aspects of both the visual and auditory elements of the map, and the relations between the procedural aspects of the various elements demonstrate many of the characteristics of digital media suggested by Manovich. Whether implemented using raster or vector graphics, the visual map elements use modular, digital media technologies based on numerical representations, and are amenable to variability in display and automation. Similarly, the auditory elements of the map are implemented either as digitally recorded audio or as synthesized compositions and each type of audio is amenable to procedural manipulation when being played as part of the audiovisual map. At the very least, the playback of audio of either type can be started and stopped as needed, the playback volume, and other playback characteristics such as stereo balance and reverb can be controlled. The playback rate can be algorithmically altered by different means for each type of audio: the tempo of a synthesized composition can be dynamically controlled and a digital audio recording can be dynamically resampled creating, in comparison to the input audio stream, either a greater or smaller number of output samples, thereby slowing down or speeding up playback respectively.

The use of visual and/or auditory elements may be more or less invariant in a specific map with the user having, for example, no more control than to cause the map to be displayed and the sounds to play (possibly through independent decisions about each). Alternatively, the display of at least some visual elements and/or the playback of at least some auditory elements of the map may be initiated or modified as a result of actions by the user. In this case, the ease with which the user understands the relations between their actions and resulting transformations of the audiovisual map will be important in determining the user's perception concerning the usefulness of the map. In particular, the sounds and sound transformations employed in a map may need to be explained through the use of visual elements or user training and may need to respect or only carefully circumvent certain expected patterns of sound usage from other, more prevalent, forms of audiovisual media. But sounds may also influence the interpretation of the visual design and thus the overall visual and auditory behaviours of the map must be considered together.[4] For the sounds, considerations must encompass what is to be heard and the ways in which the sounds are to be reproduced and manipulated.

Many types of audio recording carry implications of time and place embed-

[4]Théberge (2005) argued for sound design for cartography to be conceived as more than merely the addition of sound to maps. For discussion of the influence of sound on the interpretation of visual media, see Chapter 4.

ded within them. Recorded narrations are coded by language and dialect, accent, gender, subject matter and other characteristics that may place the narrator more or less specifically or may indeed place the narrator external to the story (Nichols 1985). Music can be understood through perspectives on genre, ethnicity, gender, or other characteristics to invoke a sense of time and place or to cite adherence to a specific canon such as Western classical and thereby reference what has by some been interpreted as universalism (for a critique of such assumed placelessness, see Leyshon et al. 1998, 3–9). Sound effects may be place and/or time specific, if sufficiently recognizable and assuming that the audience has the knowledge to decode the reference; may be merely suggestive of a type of place (e.g., a windswept plateau); or may be generic and relatively placeless. With all types of audio recording, technical manipulation of the sound, including placement and selection of microphones and the use of post-recording technology, may be able to enhance or alter a sense of the space within which the recording was actually made. And of course, recorded audio need not be thought of as statically positioned in space. Transportation technologies and the mobility of audio technologies each on their own enable motion during sound recording and create potential for thinking of a single recording as variable with respect to both time and space. Narrative reference can identify multiple places along with people, objects and other concepts within a single audio recording, thereby defining the recording in relation to a network of possible referents that again could be thought of as relevant to multiple times and places.

In addition to any spatial characteristics communicated acoustically, audio may be reprocessed in its use as part of a programmable interface such as an audiovisual map while the user works with that interface. The audio may be incorporated into the map through the use of controls that determine when or whether the audio is played, or that determine the playback characteristics of the audio such as volume or stereo balance. By combining multiple audio recordings through the use of dynamic playback controls, each audio recording could be presented as oppositional or complementary to the others, the difference needing to be deduced from the content of the audio and other cues made available through the design of the map.

Figure 3.1 represents three possible aspects or dimensions through which sounds could be used to create or reinforce meanings within audiovisual maps, including: *language and semantics*, *signification through acoustic variation*, and *linking sound with visually mapped time and space*. Each is discussed below but these potential bases of cartographic sound semantics are represented in Figure 3.1 as dimensions because different types and uses of sound offer differing possibilities and scope with which meanings can be created and communicated. Although certain sound types, such as narration, may carry very focused, encoded meanings, other types may convey general connotations, atmosphere or mood. By changing acoustic parameters controlling the playback of a sound according to potentially varying values of mapped data, additional meanings beyond those interpretable from the selection and reproduction of the sounds themselves may be communicated (if the user understands the basis of the acoustic variation). Finally, within the context of audiovisual maps, the means by which the relations between the sounds heard and the visual components of the map are presented, explained, and varied as a user works with the map may be used to clarify the interpretation of the map.

Language and Semantics. By selecting a sound to be used within an audiovisual design, there is the possibility that the sound, even if experienced independently of any visual signification on the part of the map, would carry meanings

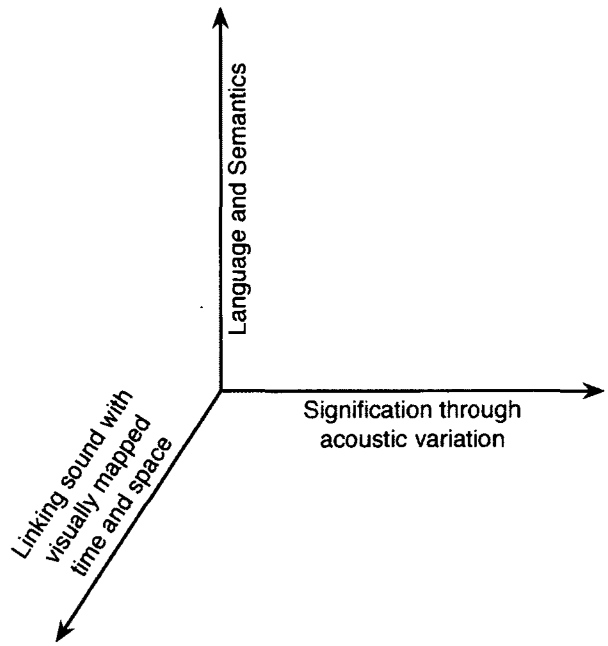


Figure 3.1: Semantic dimensions of sound in audiovisual cartography (figure concept and design: Glenn Brauen).

interpretable by users of the map. In addition to any spatial meanings or connotations conveyed by the audio as discussed above, such meaning could be created, for example, through the inclusion of recorded narration, sound effects, music, or any combination of these. Recorded narration in a language understood by a user would obviously be capable of framing a map's meaning for a user and even narration in a language a user does not understand may convey connotations (Barthes 1967, 1972). Sound effects, potentially recognizable by a user, may convey meaning through an understanding by the user of what created the sound and where or when it was recorded. Although there have been debates concerning the status of music as a semantic or a semiotic medium (for a summary, see Nöth 1990, 429–434; cf. Fiske 1990), it seems clear that possibilities exist for music to convey referential semantics through mimesis such as imitation of bird calls or rhythmic invocation of built mechanisms (e.g., trains), and through iconic reference such as national anthems or the association of certain instruments and types of events (Burkholder 2006). Quite apart from referential semantics, music can be used to create and rehearse themes or motifs within the structure of an atlas thereby introducing the possibility of establishing relations between thematic elements or sections of the atlas itself (Théberge 2005). The use of sounds, designed to adhere within or contradict audience expectations, can create meanings and shape interpretations in many ways as discussed for film and game sound in Chapter 4.

Signification through Acoustic Variation. Although coded acoustic systems such as language and music rely on 'acoustic variation,' here the term refers to the use of sound parameters such as pitch, duration, and loudness or higher-level compositional structures such as musical keys and modes[5] as abstract, modifiable characteristics of sounds. Control of these parameters can be associated with a variable so that playback of a representational sound varies according to the value of that variable and the manner in which it is linked to the acoustic parameter(s). Human-computer interaction (HCI) researchers looking for alternative representations of information for use by visually impaired users have experimented with such uses of acoustic parameters and some research has been conducted into the use of sound as a representational medium for cartography (see Chapter 4 for more discussion on each of these areas). Research concerning compositional techniques for music in indeterminate, non-linear, interactive media such as computer games could provide guidance for cartographic sound design (e.g., Guerraz and Lemordant 2008; Kaae 2008), although this research for game design is itself still very new.

[5]Théberge (2005) has argued that consideration should be given to the use of musical structures, such as melody and harmony, as possibly useful in creating acoustic representations for cybercartography rather than relying solely on lower-order acoustic properties.

Linking sound with Visually Mapped Time and Space. This dimension refers to how sounds are related to the visual components of a map and how those relationships are signified by the design and interactive control of all elements of the map application, thereby helping a user to interpret the meanings created or refined through those relations. In many cases, the methods of linking sound to the map along with the actions of a map user will determine the acoustic variations applied to a representation (the second dimension, discussed above). I refer to sounds as *linked* if they are related to a represented variable, probably spatial or temporal, and thus are used such that changes in the currently selected value for that variable modifies, in some way, the playback or effects processing of the sound.[6]

Not all uses of sound in interactive, audiovisual cartography can be thought of as related to specific components of the visual display or, in the case of a map, as varying spatially according to a geographical variable represented by that map. Sound can be used in multimedia computer applications in ways that prohibit a user from altering the audio playback (except by stopping the entire multimedia presentation). For example, music, sound effects, or didactic narrations may accompany other, usually visual, media presented to a user but may be unrelated to any graphic other than perhaps interactive *play/stop/repeat* controls. For the pur-

---

[6]The term 'linked' is derived from usage, such as that by Monmonier (1989), in which the term is used to describe visually linking multiple representations of a variable or representations of multiple variables through a set of interactive displays such that actions in one display cause visual updates across the entire set. The term is not overly visual and can easily be applied to acoustic representation. Other possibilities considered include *spatialized* and *mapped*. The term 'spatialized' indicates that the sound is somehow affected by interactions with space but the term is already in use to describe acoustic effects processing that alters playback of a sound to simulate the motion of the sound source through three-dimensional space relative to a listener. Three-dimensional spatialization could be used as the effect modifying a linked sound and I don't want to confuse the two different senses of the word. The term 'mapped' indicates that the sound is somehow affected by interactions with the map which seems acceptable, but I chose not to use this term because of the current predominance of terms like 'sound mapping' to describe the spatial organization of sounds as a *subject of* mapping rather than as a *representational method* used to create maps.

poses of this discussion, it is assumed that moving the cursor over or clicking on map areas, other graphics, or hyperlinked text does not modify or affect the playback of the sounds in any way. Sounds that are not designed to vary according to any such variable are referred to as *unlinked*. Krygier (1994) and Monmonier (1994) both proposed the use of unlinked explanatory narratives with maps. The use and effectiveness of unlinked, but synchronized, narration has also been studied for multimedia education (Mayer and Anderson 1991, 1992) as discussed in Chapter 4. Caquard et al. (2008) proposed cartographic sound designs based on an examination of sound designs for cinema and included discussion of unlinked uses of sound. Drawing on documentary traditions such as voice-over narration (Nichols 1985), unlinked sound is probably best understood as relating to the entire visual presentation currently before the user, although as with a film narrator, a sound could refer explicitly to a particular place within a map thereby narrowing the acoustic-visual relationship.

By contrast, a linked sound relates to visible element(s) the user is currently 'touching' and both the auditory-visual-haptic relation (i.e., how the audio and visuals to which the audio is related behave in response to user actions) and any semantics inherent in the audio itself may be used to understand the combined representation. Linked narration could provide place name pronunciations as the cursor is moved over symbols on a map, which could be especially useful for early learners or for use within communities that have oral-aural traditions. Within a map, linked auditory representations may be used to sound out additional variables associated with the currently selected feature or region and, as Krygier (1994) proposed, may be a good alternative to displaying multiple juxtaposed univariate maps. As a user works with different elements of the visual presentation, the sounds may change in a variety of ways: overall sound volume could be adjusted, loudness of each sound element (e.g., a narration or a specific instrument in a musical composition) could be adjusted, playback of the entire audio design or of each element could be stopped, started, or restarted from the beginning.

Careful selection of sounds for use in the map representation could shape the overall cartographic message for the user, possibly providing complementary information to the user through each sensory modality. Chion (1994, 5) argued that sound in combination with film visuals may provide *added value* by guiding a viewer to interpretations of the combined audiovisuals that in retrospect seem to have been apparent all along but, without sound, may not have materialized. Using sonification to represent a related or correlated variable to one presented visually may seem redundant when first experienced but may subtly highlight important differences between the two as a user works with a map, especially if using the same map to analyze different questions. In addition, the sounds chosen for the linked acoustically represented variable could guide a user to a preferred understanding of both that variable and the visually represented variable as well.

Harrower (2007) argued that, although combined used of auditory and visual presentation have been shown, in maps and other media, to increase overall cognitive capacity, redundancy between visual and auditory components reduces the beneficial effect of engaging both a user's hearing and vision by forcing the user to expend unnecessary effort correlating the same information across sensory modes without gaining additional information. This agrees with advice from Vetere and Howard (2000) and Kalyuga et al. (2004) that delivering essentially the same symbolic information (text) as both spoken narration and written text in educational multimedia should be avoided. Although these concerns must be considered and truly redundant information may become tiresome for a user, there are cases in which the argument against redundancy is not persuasive, at least if considered in relation to possible uses of a map. As argued by Chion (1994), it seems clear that the very idea of cross-modal redundancy, even with respect to a medium such as film that is accepted as audiovisual, is not well understood. As in the case of a map incorporating place name pronunciations there are characteristics of the spoken voice, including tone, emphasis, accent, and pacing, that simply cannot be transcribed into a written form.[7] The written text may be easier for a user to work with for some purposes and thus it may be useful to provide both but interface design, based on understandings of intended purposes for and intended users of the map must be considered. When considering other auditory representation forms such as music or sound effects, there is unlikely to be direct redundancy in the material presented using sound because graphic or textual descriptions of sound are hardly a substitute for listening. Furthermore, the potential for clarifying, enhancing, or adding value to the visual presentation may be substantial.

Harrower (2007) also questioned the purpose of spoken representation in relation to pictorial media arguing that they are not linear and thus it is not clear where to start the sonification, suggesting that acoustic representation be relegated for use with elements such as time lines which are already linear. Although the temporal structure of an auditory representation must be considered, this is only different from the temporal structure of an animation in that the user must be able to understand how the sound relates to the animation and, if the sound relates to only one region of the animated map, how that relationship is made clear to the user. That is, this is a design problem rather than a fundamental difference between our un-

[7] Monmonier (1992) discussed the use of synthesized voice to describe or 'read-out' information related to a feature to which a sound is linked. Synthesized voices have yet to be developed that compare favourably to the tone, pacing, and fluidity of a human speaker, especially a trained narrator. The design and intended use of a mapping application may determine which of recorded or synthesized voices are possible. If all of the required spoken passages are not known in advance, then voice synthesis may be the only feasible approach for the use of speech.

derstanding of sound and animated maps as temporally ordered representations. Interaction with a map seems to be quite promising as a basis for linearizing the selection of variable values to be represented acoustically: hovering a cursor over a region can be the basis of visual feedback highlighting selection of that region and can be used to modify playback or effects processing for representational audio based on the variable value at that mapped location. Designed feedback enabling a user to see, hear, and relate changes in the map as he or she works with it are important considerations. In addition, the pictorial nature of maps referenced by Harrower makes his reiteration of strictures against visible-audible redundancy which I cited earlier seem overly cautious. Combining spoken place names with visual interactive selection of mapped places, for example, could be the basis of a useful map of place name pronunciations.[8]

Finally, as discussed in Chapter 3, the emergence of the World Wide Web as the main distribution channel simultaneously for maps along with written text, music, and, possibly quite soon, movies is a recent phenomenon and the implications of these transformations for use of each media thus distributed are not well understood. It seems likely that map use will be transformed as expectations arising out of hybrid media designs and concurrent use of multiple media suggest new possibilities as have already been seen in the adoption of maps as tools for organizing real-time communications (Miller 2006). The possible uses of maps have always been subject to the expectations, interests, and understandings of cultural groups that see value in their use and have the requisite skills and resources to make use

[8]The 'spoken names' map of the *Kitikmeot Place Name Atlas* is an example of one such map. Developed as a joint project of the Kitikmeot Heritage Society and Carleton University's Geomatics and Cartographic Research Centre, early versions of this atlas used the sound subsystem to be described in Chapter 7. Web audio standards have changed since the initial releases such that, for the purposes of this atlas, simpler solutions now exist to allow sounds to be used in the map as desired and the current version no longer uses the sound subsystem developed in this thesis. The atlas is available at http://atlas.kitikmeotheritage.ca/index.html.

of maps as tools for their purposes. The intermixing of maps and other media may engender new ideas for map use, and the World Wide Web seems as likely as any channel on which some of these experimental uses will occur, but specific cultural processes and the purposes that these uses hope to fulfill (i.e., contingent expressions on the part of specific people or groups in a specific context) precede the technological developments. In social semiotic terms, such maps, within a mix of media types that continue to undergo transformation and hybridization, are *semiotic resources*:

> objects that have been drawn into the domain of social communication and that have a *theoretical* semiotic potential constituted by all their past uses and all their potential uses and an *actual* semiotic potential constituted by those past uses that are known to and considered relevant by the users of the resources and by such potential uses as might be uncovered by the users on the basis of their specific needs and interests (van Leeuwen 2005, 5, emphasis in the original).

In this research, I have followed a process of remediation (Bolter 2007; Bolter and Grusin 1999), explicitly examining existing digital map forms and audiovisual media in an effort to learn from and adapt earlier research and apply it to audiovisual cartography. In the next part of this thesis, this reuse of research and knowledge developed around other media (especially film and games) is made explicit. In addition, I discuss how audiovisual cartography can be thought of as an outgrowth of earlier research into interactive and animated cartography and I propose a functional typology of uses for sound in audiovisual cartography based on that analysis.

Part II

Audiovisual Cartography: Why and How to Use Sound?

This part of the dissertation examines why and how sound has been used in audiovisual media (film and games) and in human-computer (HCI) research. Building on this examination of sound use from other disciplines, literature concerning research into (predominantly silent) *interactive* and *animated* cartography is examined in an attempt to derive guidance from the motivations and goals behind that research. Finally, a typology of cartographic functions is proposed as the basis of a discussion of how sound could be used to complement visual design as part of audiovisual cartography in attempting to study problems similar to those studied through the use of interactive and animated mapping.

4

Functions of Sound in Audiovisual Media

This chapter examines sound design approaches used in film, games, and human-computer interaction (HCI) research and how these may or may not apply to audiovisual mapping. The existence and popularity of the first two of these types of media, and the extensive use of sound in the case of film and games in conjunction with visual design, may condition expectations among users of audiovisual cartography for how it uses sound because of the paucity of sound use to date in cartography. Audiovisual cartography could learn from results obtained in HCI research concerning the use and perception of sound as part of computer application interfaces and could potentially contribute to HCI research concerning the relations between spatial, visual, and acoustic representation.

A map is not a movie. For the purposes of this discussion, I am assuming that a map is not a game although there are many games that incorporate maps. Despite being a different form of media, designed and produced for different purposes, there are lessons for the design of sound for cartography that can be derived from the literatures on sound design for film and games, much as Caquard (2009) has cogently argued that several now-common visual metaphors in use in interactive and animated cartography were originally conceived in film production. Sound design for games, although having a much shorter history than that of film and deriving much from film sound methods while also extending and adapting those methods, shares certain characteristics with interactive cartography that make research from that domain potentially transferable.

HCI research covers a broad range of interface types, designed to address many application domains and potential user groups. In this chapter I will apply the term HCI broadly in discussing applications providing informational interfaces concerning a range of topics but narrowly by focusing on applications that incorporate sounds in ways that may provide lessons transferable to audiovisual cartography.

In first analyzing film sound design for methods and motivations that may be applicable to audiovisual cartography, I will briefly touch on three forms or tendencies in film production that have been theorized as having distinctive objectives that could influence the reasons for which sound is used and how sounds are conceived, recorded, and reproduced: *narrative film*, *cinema of attraction*, and *documentary*. Although clear examples of each tendency exist, sometimes as complete films and sometimes as scenes or objectives within films, many films are hybrids and my purpose here is to outline the objectives of sound design in the service of these tendencies more than to delineate categories of film.

Cinema, dominated by American practice, is predominantly understood as a narrative form. According to Genette (1988, 13–15), an emergent 'narrative' arises from the combination of a 'story' delineating the events described, and the act of 'narrating' that conveys those events. Because the selection of perspective for the act of storytelling (e.g., the relationship between the narrator and the events related or the mood adopted in narrating) is critical to the resulting narrative, the choices made in an act of narrating, the story told, and the resulting narrative are to some degree inseparable. Although Genette (1988) developed his definition of narrative as an analytical method to be applied to literary texts, Kozloff (1988, 42-49) adapted this definition to film narration and, in particular, to voice-over narration. She further argued that, although the applicability of the method to narrative film did not indicate an equivalence of that form and literary texts, her analysis demonstrated that problems encountered in applying the method to film could be overcome (ibid., 127). Lothe (2000, 3), in applying Genette's definition of narrative to fiction and film, argued that "[a] narrative presents a chain of events which is situated in time and space" and that the definition was applicable to fiction film on the basis that "the most important components of the definition we have given of a narrative — time, space, and causality — are central concepts in film theory as well" (ibid., 8).

According to Gunning (1986), the narrative impulse was probably not the primary motivation in the making of films during the earliest periods of cinema (prior to 1906). Instead, Gunning argued *attraction*, "exhibitionist confrontation rather than diegetic absorption" (ibid., 234), was an early motivation for cinema, drawing on vaudeville traditions, sharing its exhibition spaces, and typified by visual displays such as magic tricks and travelogues. This emphasis on visual display indicated "a conception that sees cinema less as a way of telling stories than as a way of presenting a series of views to an audience, fascinating because of their illusory power" (ibid., 230). Gunning further argued that after narrative became the predominant focus of cinema, the cinema of attraction continued in avant-garde film and at least episodically within narrative films. Scenes containing musical performance and those that rely heavily on special effects and staged stunts are examples in which the narrative progress of the film, the linkage of events through causality, may be less important than performance as spectacle. Verstraten (2009, 3–5) argued that film audiences have now become so familiar with certain plot devices that they will often fill in gaps in the narrative flow, allowing filmmakers to focus on providing entertainment without worrying about establishing complete chains of causality.

Nichols (1991, 18) argued that, in contrast to the focus on an imaginary story world as in narrative film, "[d]ocumentaries take shape around an informing logic [... that ...] requires a representation, case, or argument about the historical world." Whereas structuring a self-contained story universe, the *diegesis*, and progressing a narrative are primary in narrative film, Nichols argued, documentary is structured to maintain and propel an argument about some aspect of the world, a referent external to the film itself.

> Though shaping a plot and making a case may involve similar forms and strategies, they are also distinct. Stories occur in an imaginary universe however closely based on actual events or characters. Arguments occupy an imaginary space (they are abstract), but, in documentary film, they address or represent issues that arise in the lived, historical world. Stories, characteristically depend on plot; arguments on rhetoric. Stories must be plausible; arguments must, in addition, be persuasive. (ibid., 20)

The motivations and methods for film production and design often do not separate cleanly along the lines of the division briefly outlined above. However, for my purposes these tendencies can be argued to provide motivations for sound design choices that may need to be made when creating an audiovisual map. Depending on design objectives and the audience intended for an audiovisual map application, creating a sense of narrative within it may be useful in providing explanation concerning those objectives. It may be equally important in some cases to provide enjoyable visual and acoustic material as part of the application to draw users in and to help them to engage with the subject matter. Considering the frequent use of maps as tools for understanding and communicating spatial information concerning a variety of aspects related to the historical human world, it may be worth considering how the designed visual and acoustic components of a map relate to the objectives of the map and in establishing (or not) that the information in the map is, as Nichols (1991, ix, emphasis in the original) argued for documentary film, "*evidence from* the world." Whereas he argued that a film, documentary or otherwise, is always the result of conscious planning, framing, editing, and production choices there is an expectation among documentary viewers that what is shown and heard in documentary is similar to that which would have happened had a film crew not been present. Similarly, maps, despite the presence of authorial influence, are often accepted as authoritative and an assessment of how the incorporation of sound, or types of sounds, may support or weaken that authority may be appropriate.

These tendencies may be compounding in guiding a specific choice such as if a sound recording is both useful because it establishes narrative elements within a map or atlas and because it explicitly references the historical world, possibly by reproducing recognizable voices or the sounds of events. By contrast, these tendencies may sometime be contradictory and, as in visual cartographic design, decisions concerning trade-offs may be necessary.

Although attraction and documentary impulses seem highly compatible with audiovisual cartography, the narrative tendency and how it applies requires some clarification. Manovich (2001, 218–228) argued that digital media artefacts tend to follow two different types of organizational logics: *narrative*, a linear explanation of causality, and *database logic*, the storage and indexing of data collections so as to permit alternative access and reordering of the data. Manovich argued that database logic is common in new media forms such as virtual museums, digital encyclopedias, and are blended with narrative to create hybrid media forms such as computer games. Adapting the definition of narrative to include digital media artefacts, Manovich (2001, 227) argued, following (Bal 1985, 8), that a narrative "should contain both an actor and a narrator; [...] and its 'contents' should be 'a series of connected events caused or experienced by actors.'" In the case of a computer game, for example, Manovich equates the player and the actor. Unstated but implied in this formulation is the equivalence of the narrator and the author or artist that created the game. Manovich is clear that the simple possibility of browsing database records in an arbitrary or random order does not create a narrative. Rather, a logic establishing causal relationships between events through which actors are propelled as a result of their actions and decisions creates a narrative, possibly one such chain through multiple paths available in hypernarrative.

Returning to the discussion of maps, Monmonier (1994) argued that there are two main metaphors guiding the development of animated and interactive cartography: *navigation* and *narration*. Navigation, argued Monmonier, is apparent in software interfaces that allow a user to locate images and other information about places, regions, and spatial relationships. The user directs the system to narrow, broaden, or adjust a visual display to satisfy informational desires. Narration, by contrast argued Monmonier, is apparent in system functions that operate by playing out sequential displays of information based on predefined scripts. "The presentation is narrative rather than navigational because the user is now a comparatively passive viewer, who watches while the system [...] controls the sequence of scenes" (ibid., 202). According to this analysis, he divides the uses of digital mapping systems into navigational or narrative based not on the underlying information-seeking actions of a user but based instead on which of the user or the computer is more in control of the display. It may be most appropriate to think of this as a continuum on which different systems or indeed different sequences of interaction of a user with a single system may show variable influence of each of the navigation and narration metaphors. However, Monmonier's analysis highlights the temporal context involved in the use of interactive mapping systems. There is a sequencing, possibly conceived as a 'dialogue,'[1] as a user directs the navigation, requests and receives feedback in forms exhibiting characteristics of narrative to a greater or lesser degree on where the navigation has thus far led him or her in a geographic investigation, and plans further steps for the investigation.

Although Monmonier's use of the term 'narrative', deriving as it does from his discussion of 'narration', is different than the earlier definitions I have been using in this discussion, there are parallels between his discussion and that of Manovich. Continuing the equivalence between actor and now map user (formerly player), the map's construction at least partially implies a logic similar to that of the narrator's Manovich declares so important in determining whether or not a digital media artefact is in fact a narrative construct. Rather than a random assemblage of database records, a map provides organizational, thematic and spatial logic guiding the interactive dialogue between user and map. The syntactic structure of the map and linear textual narrative are very different but this has not prevented researchers from arguing that even static paper maps have a syntactic structure.[2] Furthermore, Manovich in discussing games exemplifying digital media narratives discusses the games *Doom* and *Myst* at length, emphasizing the importance

[1]Monmonier in fact uses the term 'dialogue', comparing the interactive user-computer interchange to that of a diner and a waiter in a restaurant.

[2]Ratajski (1975, quoted in Head 1991, 243–244), while working on the syntax of compound map symbols, argued that "the syntactic structure of the object language of cartography [is] a reflection of a structure of concrete space" (*ibid.*, 10) and argued further that a "map is never read 'from cover to cover'" (*ibid.*, 16) but that map reading is rather a sequence of concentrations upon smaller map segments and comparisons of those segments to the map reader's mental conceptions of the mapped space.

of navigating three-dimensional space in both advancing the narrative and in the aesthetic of these games (ibid., 244–247). Although there may be more recognizable 'events' occurring while navigating these game worlds than while browsing, interacting with, updating, and consulting an interactive map, depending on the subject of the map, the difference is possibly no more than a matter of degree. For example, Charles Joseph Minard's *Carte figurative des pertes successives en hommes de l'Armee Français dans la campagne de Russe 1812–1813*, a static map showing loss of life by the French army during Napoleon's Russian campaign, is tightly packed with events if only a person is ready to interrogate and re-imagine the meaning of the symbols. As with other media forms, it may be appropriate to think of maps and digital maps on a continuum, some exhibiting characteristics of narrative more strongly than others.

Manovich's database metaphor is also strong in geographic information systems (GIS) and maps, with some definitions of GIS explicitly including reference to the role of the system as a geospatial data store. This data storage role was historically filled by paper maps before the introduction of digital mapping systems emphasized the role of the map as an interface to the underlying database. But the strength of the database metaphor in these technologies does not necessarily weaken the narrative metaphor of map applications created using them. The application of cartographic selection, generalization, framing, and general design choices made in creating a map as one of possibly many interfaces to an underlying database establishes the narrator's logic demanded by Manovich's definition. A user, through the mediation of the map as interface, may be far more aware of the map's subject matter and the logic of the relations represented than of the underlying database technologies. As discussed above, the characteristics of that subject matter and its presentation may finally have a more pronounced influence on a user's understanding of the digital map and its tendencies toward narrative, documentary, and/or attraction.

Within any particular audiovisual mapping project, consideration of a range of uses of sound may be worthwhile at the level of a single digital map or at the level of an atlas comprised of multiple digital maps. Using Manovich's definition, a digital map on its own may to varying degrees demonstrate a narrative structure. The insertion of a linear narration related to the entire map or a set of narrations, each possibly related to a portion of a map or certain features on a map, may enhance the map's narrative structure. Alternatively, a map used to organize a set of audio recordings spatially may highlight the database organization and logic of the map, with or without adding a narrative component to the map depending on the subject matter of both the map and the audio recordings.

Wood (1987) argued that an atlas always contains a narrative structure relating maps to other textual and visual material and embedded in the underlying logic dictating the ordering of the maps in the atlas. Again using Manovich's definition of narrative, a user's interrogation of an atlas and the logic implied by its construction would seem to fit, but it may depend on the subject matter of the atlas and whether or not (or too what degree) that is interpreted to communicate events. The logic of ordering by which an atlas is constructed may also appropriately be interpreted as analogous to the documentary tendency described by Nichols's (1991) definition as an argument about the world and as a distillation of information and evidence concerning the world. Therefore, the use of audio within an atlas could emphasize or enhance narrative structure, or could provide an additional form of evidence in support of an atlas's documentary arguments organized through database technologies.[3]

---

[3]Depending on the objectives of the producers of an atlas, the possibilities available through

4.1 Functions of Film Sound

From the earliest exhibitions of recorded moving images, the practice of accompanying imagery with sound was almost ubiquitous (Altman 1992a; Berg 1976; Cavalcanti 1939). Silent film exhibitions made extensive use of music, often by employing live musicians but also using non-synchronized recordings when such technologies became available and when an exhibitor's business did not warrant or could not sustain the greater expense of live music. The exhibition of film with no musical or acoustic accompaniment was found to be poorly accepted by early audiences (Berg 1976, 24). Although experiments were conducted with lecturers providing narration and actors voicing dialogue, this did not achieve as high a level of acceptance as did the use of background music, sound effects and mimetic music although the latter two were, after an initial period of novelty and fairly broad use, restricted almost exclusively to comedies (Berg 1976, 93–96). By contrast, Altman (1992a, 35–37) argued that many accounts of early film sound accompaniment ignored the diversity, sophistication, and popularity of some early practices, citing the number of commercial ventures developing systems and the variety of touring acts using live synchronized sound effects and human voice accompaniments.

The transition from silent to synchronous sound films, in which the distributed print carried a soundtrack as well as the moving imagery rather than leaving decisions concerning acoustic accompaniment to exhibition houses, began during the

varying the narrative and database metaphors at play in the form of a digital atlas, and obviously through selection of evidence to be included, could result in a range of different outcomes. Emphasizing the database character of an atlas even if approached ethically does not necessarily reduce the play of the the narrator's logic in the resulting artefact. Alternatively, opening the possibility for community contribution to the evidence to be included in an atlas through the use of those same database technologies creates the possibility for admitting multiple perspectives within an atlas (Brauen et al. 2011; Taylor 2005c).

late 1920s and extended into the 1930s (Gomery 1976, 2005; Stewart 1980). During this time, competing interests and traditions promoting candidates for inclusion in the acoustic accompaniment of films staked out claims for accommodation on the soundtrack. Audience interest in synchronized speech, partly because of its novelty, created an early focus on its use. Traditions of musical accompaniment and sound effects developed for silent films were maintained (Cavalcanti 1939).

In the following sections, I briefly discuss each of the primary soundtrack components: music, voice, and sound effects. This is not to ignore the importance of the relationships between these components as is highlighted by the conception and common practice of dynamically mixing those components to create a final soundtrack. Of course, classifying sounds cleanly into categories in the first place is problematic because acoustic properties such as rhythm or pitch underlie, to a greater or lesser degree, sounds in each of these categories and can be emphasized or understated as desired for compositional or dramatic purposes. Along with each type of sound, additional consideration is given to the spatial effects of sound, because of the particular pertinence of space in cartography, and the use of silence, because of the marked difference between film and cartography in its use.

4.1.1 Music

Music is sometimes an intrinsic part of a film's story, as when exhibiting filmed ballet, opera, or the biography of a composer. In these cases, the subject music would ideally accompany the film. Manvell et al. (1975, 88) contrasted these *realistic* musical accompaniments with *functional* accompaniments in which the music "'points', underlines, links, emphasizes, or interprets the action, becoming part of the dramatic pattern of the film's structure." However, Berg (1976, 43) argued that it was not initially apparent that music should play such a functional role in relation to the image and approximately 15 years passed between the initial uses of sound accompaniments for silent film and the realization among musicians and exhibitors that "the musical accompaniment could and should relate to the moving images."

As a narrative support, Berg argued that film music fulfilled the following (overlapping) functions (ibid., 170–216): identifying and characterizing roles (e.g., 'hero', 'villain') or recurrent situations and ideas (e.g., 'battle') through the use of a repeated musical theme or *leitmotif* (ibid., 171); establishing atmosphere by relating a sense of time and place and a cultural context within which the story action takes place (ibid., 182); reflecting actions or types of actions through a range of techniques including establishing a mood appropriate for a particular activity (e.g., robbery) or synchronizing particular sounds to the rhythm of visual activity (e.g., a character's footsteps) which came to be known as 'Mickey-Mousing' (ibid., 190–191, 199); and establishing moods or emotion (ibid., 197). Despite discussing music in the context of films that included synchronized soundtracks, Manvell et al. (1975, 88-89) present a very similar list of categories as a framework for their discussion of functional film music. In addition to the use of music as a functional support for a film's narrative structure, Manvell et al. also argued that the use of music "allows the producer to linger in certain scenes to an extent he would not dare otherwise" thereby helping to create time within which that structure can develop.

For Gorbman (2003), music's inclusion as an element which exists outside of the film's narrative space (i.e., music that cannot be heard by the characters and is thus referred to as *non-diegetic*) must have persisted through the transformations from silent to narrative films with synchronous sound because it satisfies certain needs in the exhibition of films.

> Music has persisted as an integral part of the sound film because it accomplishes so many things at once. Its freedom from the explicitness of language or photographic images, its useful denotative and expressive values easily comprehended by listeners raised in the nineteenth-century orchestral tradition, its malleability, its spatial, rhythmic, and temporal values, give it a special and complex status in the narrative film experience. [. . .] It bonds: shot to shot, narrative event to meaning, spectator to narrative, spectator to audience (ibid., 39).

McClary (1991) argued that music carries cultural constructions of gender and sexuality and actively participates in negotiations and portrayals of gender, rather than just passively reflecting already existing definitions. Furthermore, she argued that there has been a tendency to treat music as a gendered discourse: "music's association with the body (in dance or for sensuous pleasure) and with subjectivity has led to its being relegated in many historical periods to what was understood as a 'feminine' realm" (ibid., 17).

4.1.2 Voice

Despite the early emphasis on music to accompany silent film and the continued use of music once the narrative film industry had transitioned to synchronized sound, recorded voices in synchronized sound cinema are considered the most important part of the soundtrack in advancing narrative and rendering it intelligible. Kozloff (2000, 6) argued that in American films "what the characters say, exactly how they say it, and how the dialogue is integrated with the rest of the cinematic techniques are crucial to our experience and understanding of every film since the coming of sound." Written and staged to appear as if intended to communicate information among the characters, film dialogue functions to bring the audience into the story and to clarify and rehearse identities, intentions, and expectations as necessary to keep the audience tuned in and interested in the narrative (ibid., 18–19, 39).

The representation of characters speaking and conversing within the diegesis is only one use of recorded voice in film. Voice-over narration in which a voice is heard but no speaker is visible is often used to provide needed information to the audience to assist in comprehensibility and to help set atmosphere for a film (Kozloff 1988; Nichols 1985). Audience members determine whether speech from an invisible speaker is a voice-off (i.e., a character within the diegesis but out of frame) or a narrator through learned appreciation that includes what is being said, how it is said and recorded, and whether or not other characters in frame react audibly and/or visibly to what is said.

First-person narration can provide information and background concerning events and chronologies; can increase audience identification with and empathy for characters; can emphasize the subjective aspects of a story because of the direct address by the narrator to the audience; and can reveal the thoughts and feelings of a character to the audience while shielding this information from others in the story (Kozloff 1988, 41). Third-person narration by a speaker that never appears before the camera can convey complex expositional information and may literally declare the perspective of the filmmakers.

For documentary with its concern for arguments about the historical world, Nichols (1991, 20–21) argued the ability of speech in a variety of forms to efficiently communicate and shape arguments as the core of a film heightens the importance of the voice and it is used extensively in the form of voice-over narration and in the form of commentary and interview response by subjects and participants. Nichols (1991) argued that, at the time he was writing, there had been four predominant modes of documentary filmmaking, characterized by the narrative devices employed by the filmmakers in structuring their films' arguments: *expository*, *observational*, *interactive*, and *reflexive*. Each different mode, he argued, created a different type of claim to providing the viewer with access to aspects of the historical world about which the film was concerned. The expository mode uses direct address to explain background information, implications, and often solutions with visual montage used to support the argument. The expository mode establishes an authoritative tone concerning the events or situations depicted and continues to be used regularly, including on network television news broadcasts. Each of the other modes discussed by Nichols repositions and renegotiates authorial control and power in the representation of people, places, and events depicted, although as he argued final control always rests with those in control of the camera, microphones, and editing processes.

Alternative modes were enabled by more mobile technologies that made it possible to record images with synchronized sound, and thereby forego the use of post-production voice-over narration in expressing a documentary argument. In observational mode, a desire to stress "the nonintervention of the filmmaker" (ibid., 38; also Ruoff 1992) and ethical concerns about the representation of people and the world were the motivation for developing an alternative to expository mode. As opposed to the editing in expository mode, predominantly concerned with supporting a documentary's argument, Nichols argued that editing in observational mode "serves mainly to sustain the spatial and temporal continuity of observation" (Nichols 1991, 40). But observational mode documentary, with its distaste for expositional mechanisms such as voice-over, inter-titles, and explanatory text "limited the filmmaker to the present moment and required a disciplined detachment from the events" (ibid., 33).

By contrast, interactive mode emphasizes verbal exchange and dialogue with recruited participants, the filmmaker often heard off-screen asking questions and making comments, and the use of editing and visual imagery mainly to corroborate or question the viewpoints expressed by participants. In order to engage with people and events directly, the filmmakers adopting interactive modes chose to eschew the careful detachment of observational mode: "looking, listening, and speaking" (ibid. 44).

Reflexive mode documentary, Nichols argued, although using techniques similar to interactive filmmaking, focuses more on the relationship between the filmmaker and the viewer than on that between filmmaker and subject. "It stresses the deformative intervention of the cinematic apparatus in the process of representation" (ibid., 61).

Regardless of the positioning of the narrator with respect to a fictional narrative, either as a character or as exterior and omniscient, Kozloff (1988, 98) argued that "all narration requires knowledge, reveals biases, connotes power." The positioning and claims to knowledge made by those speaking in documentary also demonstrate a range of authority in relation to the documentary argument and the people and events depicted. Voice-over narration in documentary still claims an authoritative perspective, audible off-screen direction can reveal an editorial perspective guiding respondents, while on-screen speakers are assumed to speak from a more contextual perspective as participants.

4.1.3 Sound Effects

Within film, the term 'sound effects' refers to the synchronization of sounds with narrative events, whether on-screen or off. The sounds produced could be intended to increase the sense of realism of a depicted scene or to underscore and mimic the rhythms of visible movement, as was already discussed above with respect to music. Even before the introduction of synchronized sound on film, sound effects were being produced as part of the musical accompaniment of film, drawing on vaudeville traditions, initially through the use of percussion instruments and horns and eventually using organs with built-in sound effects (Berg 1976, 191). The distinction between sound effects and music in film is often not clear because care is taken to produce sound effects that work as part of the sound design along with the music.

Although sounds synchronized to actions depicted on film are intended to be interpreted by the audience as coming from those actions, film exhibition developed techniques to substitute sounds for direct recordings of the action before the technological capabilities to capture and reproduce those sounds were available. Even after synchronized sound was introduced to the film industry, post-production techniques to re-record and produce a final soundtrack including sound effects to accompany the visual action continued and developed. Direct recording of a scene to capture dialogue, sound effects important to the narrative, and the music used on the film would create an impossibly complex and expensive production regimen. Instead it is taken for granted that controlling sound during filming of scenes may not be possible and post-filming re-recording of at least some elements is planned. Furthermore the sounds that are included are often creative fabrications substituted for the sounds produced by the action depicted because: a 'hyper-real' sound, exaggeratedly loud or dramatic, can be more clearly heard (Holman 2002, xviii); the original sound may be difficult to record clearly or expensive to edit into the film as frequently as needed (e.g., footsteps) (LoBrutto 1994, 61–68); or the sound is intended to underscore a dramatic interpretation of a scene rather than represent an existing reality (Murch 2005, 15–17).

In addition to interactions between sound effects and music based on developed film traditions and aesthetics, there are important relations between voice and sound effects. Whereas dialogue can efficiently identify, name, and explain narrative or documentary background and motivations, sound effects can efficiently ground the narrative in a recognizable auditory world of events and atmospheres. Mott (1993, 7) argued that early radio dramas suffered from stilted dialogue conventions before abilities were developed that enabled producers to create sufficiently realistic sound effects that allowed dialogue to refer to sounds such as a car arriving without having to explain things that were both invisible and inaudible to the audience. Within film narrative, an off-screen sound effect, allowing the visual image to linger on a character's response was found to be an effective alternative to either the same kind of dialogue problems cited by Mott or the need to visually show both an event and its emotional response (Cavalcanti 1939).

As with voice, producers of documentary have sometimes developed their own conventions concerning the recording and production of sound effects. Ruoff (1992) argued that observational documentaries, attempting to demonstrate fidelity to the events depicted and avoidance of obvious narrative devices, would retain the location sound recordings from sometimes extended film takes under less than ideal recording conditions. As a result, in comparison to narrative fiction films, these tend to include a broader range of sounds in the final film, the dialogue tends to be recorded with higher levels of ambient sound, and comprehensibility is often not as good. However, Ruoff further argued that documentaries exhibiting predominantly observational mode characteristics sometimes adopt strategies as needed to balance comprehensibility with accepted genre practices. If in the process, these documentaries mix techniques such as post-production re-recording of environmental sounds and voice-over narrations with observational filming, he argued, the results do not inevitably demonstrate "that strictures for making documentaries were violated, or that audiences were necessarily deceived, but rather that all films are constructions, meaningful assertions about the world made by directors and those with whom they collaborate" (ibid., 234).

4.1.4 Accommodating the Mix: Comprehensibility

As discussed above, none of the sound types used in a film is independent and ultimately all needed sounds must be accommodated as a mix. In early synchronized sound productions, components of the soundtrack tended to be produced without regard for the auditory spectrum other components would require with, for example, music being produced in isolation from the dialogue and vice versa. According to Altman et al. (2000, 343), the resulting soundtracks were often difficult to understand because insufficient consideration was given to the overall mix and those responsible for each of the component sounds, orchestra leader and actors, would work according to their previous experiences with, for example, silent film and theatre, respectively. But over time, Altman et al. argued, greater integration of the various soundtrack components was achieved through a conceptual recognition of the need and the technological means required to create a mix of dialogue, music, and sound effects, varying at least the volume levels of each throughout the running time of a film to ensure narrative clarity.

The predominant approach, generally accepted with little reservation despite the debates during the years of the transition to synchronized sound (Altman 1992c), was summarized by Murch (2005, 14) as "[t]he human voice must be understood clearly in almost all circumstances, whether it is singing in an opera or dialogue in a film." Although the predominant method of recording dialogue for American narrative films has been to use close microphone placement to reduce reverberation and maximize clarity, techniques vary as deemed appropriate to a particular story. Close miking can imply intimacy and access to a character's inner thoughts (Ondaatje 2002, 64–65). Distant miking can allow reverberations to develop in the interactions of a character's voice with the space and architecture around the character and can be used to reinforce the place of the narrative and to allow louder pronouncements by the character to take on certain emotional tones such as anger and passion (Klimek 1992, 207–208). As Klimek argued, the linear narrative structure of film allows a production to use a "wise microphone" (often along with anticipatory camera movements and selection), switching and blending quickly between microphone placements and other post-production choices as required.

Concerns over film comprehensibility may arise due to a variety of issues involving the visual composition and the soundtrack. But within the soundtrack, concerns about comprehensibility most often centre on whether or not dialogue and narration can be clearly understood, leading to "the subordination of all other considerations to ensuring that the spectator can hear the words fully and well" (Kozloff 2000, 120). According to Altman (1992c, 60), the desire for intelligibility of dialogue forces a substitution of a filmic model of reality, traced by him to theatre traditions, in the place of "the everyday life model." Degradation of dialogue and narration comprehensibility could be caused by script issues such as complexity of language used, dialects, and overall writing quality or by performance issues such as character accent, enunciation, interference from overlapping sounds or speakers, and recording levels or characteristics.

Despite the desire for dialogue clarity, it is sometimes sacrificed to use voices as part of the sonic atmosphere of a film. Kozloff (2000, 189–190, 215–216) argued that in certain genres intentional use of techniques known to reduce comprehensibility serves narrative purposes. For example, dialect, slang, accents, and overlapping dialogue are all used in gangster films to signal group membership and coherence. Chion (1992, 105) referred to this type of dialogue usage as *emanation speech*, distinguished from more common *theatrical speech*, and argued that it is not usually central to the narrative's action and thus the viewer's understanding of it is not as important as its effect as "an aspect of [the characters] themselves, like their silhouette: significant but not essential." Within a carefully constructed soundtrack, characters can switch between theatrical and emanation speech quickly to achieve atmospheric goals such as those described above. Frequent or prolonged use of emanation speech could very likely detract from a viewer's sense that they are understanding the film.

4.1.5 Implications for Audiovisual Cartography

In the development of film production and distribution methods and in fostering an audience for the industry, debates occurred concerning appropriate methods of recording sound for film and the appropriate objectives of sound reproduction for which the methods should strive: adherence to acoustic realism or narrative intelligibility. During the first decade after the introduction to film of synchronized sound, the methods that became predominant, although not ubiquitous, were those that supported narrative intelligibility. But rather than clarity being maintained as a constant top-of-mind goal, film producers balanced elements of the soundtrack over the linear progression of the narrative. Thus they could sometimes favour those sounds that were important for the audience's understanding of events, often narration and dialogue, while at other times allowing the environmental sounds to establish setting, assisting in the willing suspension of disbelief, or perhaps the music to assist in setting an appropriate mood. The approach to the design of the soundtrack became temporally dynamic.

In addition to explaining some of the complexity with which film producers approach the design of soundtracks, a reading of histories concerning this debate firmly establishes that film production is an endeavour in communicative design rather than an attempt to capture an already existing reality. There is no essential form of cinema and many possibilities for it remain to be explored (Altman 1992b; Carroll 1988). Sound design for film is part of this process and debates concerning the writing, design, and editing of sound for film highlight the status of film as a constructed medium.

Sound and imagery (static or moving) together create meaning and emotion, not necessarily realistically but according to conventions developed (and continuing to develop) over time and contingently embraced by producers and audience members alike. Preexisting models for music, voice, and sound effects were tried, adapted, and development continues. The possibilities for sound or imagery to reinforce, complement, or contradict the other in the service of narrative are nearly limitless and beyond cataloguing.[4]

A map is often a representation of some characteristics of spatial reality or a metaphoric spatialization of other information. In either case, like all other media it is never a mere reproduction of reality. Designed with communicative intentions, films and maps both distill complexity into a form intended to clarify characteris-

[4]Kozloff's (2000, 189–190, 215–216) argument that intentionally poor dialogue clarity is used in certain genres as a form of deliberate signification would seem to indicate the impossibility of fully developing such a catalogue. Film's self-referentiality, in this case in the expectation by audiences that filmic reality is understandable, allows producers to create semantic meaning by playing against that expectation and undercutting, at least temporarily, the clarity of the soundtrack.

tics of a topic according to the priorities of the producers. Issues of intelligibility, expressed in terms such as 'usability,' 'usefulness,' and 'effectiveness,' have also guided much research into map use and design. Among most map users, claims to realism are unlikely to trump clarity in representation. If a trade-off between the two goals seems necessary, a user's insistence upon realism would (or at least *should*) probably mean they'd be unlikely to find the map a suitable source of information in the first place.[5] In adapting cartographic design for the inclusion of sound, goals of intelligibility, both in the combination of sounds into a map soundtrack and in the audiovisual interplay of the overall design, must be a high priority. Until sound is a common element in map designs, there are no possibilities for playing against map users' expectations in the design of sound. Unintelligible sound may well be interpreted as simply that.

Although geographical methodologies have been challenged in recent decades to accommodate alternatives to historically dominant visualities, use of sound in cartography is still rarely researched. When used at all, sound tends to be considered in highly constrained manners, most frequently used to reproduce or represent a realistic aspect of the location with which a recording has been associated (i.e., a reproduced soundscape that may be, for example, historical, natural, or cultural), or less often used strictly as an abstract representation of data. These methods may align well with the documentary impulse in constructing a spatial argument and in using sounds as part of the evidence in support of such an argument, depending on the subject matter of the map. The first may be more obviously similar to the observational mode of documentary in representing location sound as evidence from the world but, as in the case of film, that will depend on many

[5]But I am reminded that I have had to tell people "no, you will not be able to see the colour of my neighbour's new roof just by looking at Google maps."

factors including what is recorded, how it is recorded, and any narrative structures with which a user engages while using the map. Such narrative structures could consist of text or narrations providing background for the map or atlas, related imagery, or additional acoustic material such as music linking this map to other material.

Depending on the narrative within which a map or atlas is embedded, sound elements added to the map may be part of and enhance the narrative. Spatially organized stories, with playback initiated through a map interface could provide contextualized perspectives, with identity and character traits of the speaker possibly communicated through spoken information and vocal traits such as language, accent, tone, and pacing along with visual indications on the map identifying whom is speaking and from where. User interpretation of such stories can be influenced by many aspects of the atlas design but again lessons from film may be germane. Located contextualized stories such as those just described may still be interpreted as being subject to selection and editing choices, but this interpretation, if of an online web atlas, could also be moderated if the atlas interface is designed to collect and make available additional stories in addition to broadcasting existing stories. If appropriate for intended audience, subject matter, and technologies, consideration could be given to incorporating a narrating character into an atlas, with or without a visual representation (i.e., possibly following radio narrative techniques as much as film) or indeed varying the character's representation throughout different sections of the atlas.

Turning to a more specific use of sound with respect to film and maps, Rice and Rice (2001) studied the musical scores accompanying animated maps in a sample of films, predominantly Hollywood fiction and animated films, created over a fifty-year time span. They adapted Karlin's (1994) classification of the ways in which music can play through a film's drama (see also, Karlin and Wright 2004, 154–164), categorizing each film in their sample as one in which the music: *hits the action*, *phrases the action*, or *plays through the action*. With respect to a map, playing through indicates that the score does not noticeably respond to the presence of the map or acknowledge any action related to the map. By contrast hitting or phrasing the action both respond to the presence of the map or the action surrounding it, the difference between the two being a matter of degree. Whereas phrasing the action comments only on significant moments requiring an emotional comment (e.g., arrival, departure), hitting accents every action.

Rice and Rice argued that all of the above approaches had been used with at least some of the map animations in their sample (playing through: 48%; hitting or phrasing: 44%) and that the music also frequently helped to establish the setting of the narrative using geographical or cultural references and used leitmotif to reference characters and situations relevant to the use of the map within the narrative.

4.1.5.1 Space

Levin (1984, 62) argued that "film music supplied the spatiality implied but not realised by the animation of the image." Whereas sound has been theorized as very important in establishing the diegesis as a three dimensional space into which the audience has a privileged viewpoint from which to watch and listen as the narrative unfolds, the paucity of sound with most maps has meant that the possibility of such a relationship has not been examined.

Doane (1985) argued that in the situation of viewing a film, spectators are confronted with three spaces: the diegesis, unbounded except by the scope of the story; the two-dimensional space of the screen, always limited by the current framing of the image; and the acoustic space of the theatre, surrounding the audience. The relationship between a voice heard in a film and the positioning of the speaker, as visible within the frame, as off-screen but within the diegesis and thus audible to characters in scene as well as to spectators, or as outside the diegesis, creates complex systems of signification that, for Doane, bring the spectator and characters into overlapping and topologically connected spaces "in the service of a representational illusion" (Doane 1985, 174).

The characteristics of a sound as perceived by a listener are shaped in part by the acoustic qualities of the space within which that sound is created (Belazs 1945). When recorded, some character of the space within which a sound was produced and recorded is retained and this carriage by sound of the qualities of its production space has been used in film to invoke a sense of a story's space for the audience.[6]

As already discussed for voice, this interaction of sound with the space in which it is recorded can be used for dramatic effect. Similarly, the recording of environmental sounds can be used to invoke a sense of place in which a scene is unfolding or, in documentary, as part of the context for the arguments and proposals to be presented. Altman (1992c) argued that, in conjunction with debates about the proper recording of dialogue, the objectives and methods for representing sounds occurring within the diegesis were also contested. But, he argued, by the end of the 1930s sound effects were often reproduced appropriate to the "point-of-audition" of characters within the diegesis that were intended to hear them (ibid., 60; cf. Maxfield 1930). Practice was and is not homogeneous. Some narrative films produced

[6]Lastra (1992) argued that it may be more accurate to discuss all sound recordings as *representations* rather than *reproductions* because of the specificity of the conditions of sound production and capture in the recording of any acoustic event and because of similar specificity in the replaying of a previously recorded acoustic event.

even after this period continued to use sound recording models that attempted to place the vision and hearing of viewers in a consistent relationship with the scene or indeed switched between models stressing spatial fidelity and models designed for narrative clarity within a single film, depending on the objectives of the production (for a discussion of these issues with respect to several film adaptations of Shakespeare, see Klimek 1992). Soundtrack music, when not intended to be heard or interpreted to be heard within the diegesis occupies a separate space from dialogue and sound effects, psychologically as well as practically (Gorbman 1980).

Considering the importance of space as a focus of geographical and cartographic study, the contrast between the conceptualization of the role played by sound in the constitution of space in film compared to the use of sound in maps would be remarkable, if not for the differences between the historical development of each media. Limiting the discussion mainly to the audiovisual design of orthographic map views, as I am in this thesis to clarify and delimit the discussion, the role that sound should play in spatializing a region represented in a map is unclear but deserves consideration and further study. The visual projection of mapped space, often based on an artificial and unobtainable point of view, suggests that realism in acoustic viewpoint would be no more useful or achievable an ideal against which spatializing sound designs should be judged than it has become in film. Although the debates concerning the methods for representing space through the use of sounds in film indicate that there is no obvious model to be adopted in representing the topologically connected spaces of film (Doane 1985), it is clear that a map may provide a very different conception of represented space as a starting point. As in the case of attempting to amalgamate lived conceptions of space with those modelled in geographic information systems and maps, predominantly understood as visual and geometric (Peuquet 2002), possibilities for accommodating representations of lived acoustic space may well be contextual and difficult to generalize.

Quite apart from invoking space, sound could be used to signify topological relations within space through technologies that enable the spatialization of sound sources in relation to a listener, cueing them, possibly after learning that this acoustic motion is intended and significant, to examine the mapped direction from which they perceive the sound is coming.

4.1.5.2 Silence

Complete silence, an absence of any recorded music, voices or ambient noises in a film soundtrack, is very rare and is usually considered to be undesirable (Figgis 2003; Théberge 2008). As stated earlier, prior to the introduction of synchronous sound, exhibition of films without any acoustic accompaniment was not well accepted. Even during a film's quietest moments, visuals are almost always accompanied by a "room tone" or ambient soundscape at the very least. Figgis (2003) argued that complete silence, when used, tends to powerfully underscore the narrative impact of the accompanying scene. But, as Murch (2003) argued, the perception of silence by film audiences only becomes meaningful, and thus potentially dramatically powerful, because of their expectations for the almost ubiquitous presence of sound.

Théberge (2008) argued that a discussion of silence in film must begin with "relational silences" understood as periods in which some of the elements of the soundtrack (music, dialogue, or ambient sound) are silenced relative to the others but also importantly relative to expectations created for that element by earlier usage, possibly within the film itself or within a broader context such as genre (see also Chion 1994, 57). Such relational silences, argued Théberge, have been used within film for a variety of dramatic purposes including (very briefly and selecting only a few of the many possible purposes): bringing the audience into the psychology of a character; compressing space and/or time during narrative transitions allowing plot to move forward quickly; or to signify important transitions within the narrative such as a death or the exposure of information critical to the story.[7]

Apart from dramatic purposes, sound editing may be used to create functional but transitory silence of some soundtrack elements to maintain narrative intelligibility. Most often this is done to ensure that dialogue is clearly audible and such that the audience is unlikely to notice the removal of the potentially overlapping sounds (e.g., music, ambient soundscape, or additional dialogue), focusing as they are expected to on the dialogue (Murch 2005). For dramatic purposes, this is not considered to be a relational silence because the intent is to support intelligibility without the audience noticing the absent sounds.

As has been discussed earlier (see Chapters 1 and 3), absolute silence is the default accompaniment for maps although it is not at all clear that the environment within which maps are used is normally silent. Differences in the status of sound and silence between film and cartography can probably be explained extensively in terms of the differences in the economic, cultural, and historic processes out of which they have developed: the silence of maps being a straightforward extension of historically silent print media and the acoustics of film being an extension of some of the predecessors of cinema. Whereas I agree with Théberge (2005) that silence could be used as a meaningful element within the design of cybercartographic atlases, an important element in establishing silence as a meaningful map

[7]See also Berg (1976, 199–201) concerning the use of silence by musicians in the silent film era to deepen the emotional impact of scenes depicting unexpected or highly emotional events, such as death, or to accompany representations of, for example, reverence, bravery, or prayer.

or atlas component would be the expectation for sound to be used. Without such an expectation, there will be no meaning in silence.

4.2 Functions of Game Sound

Games that incorporate sound as part of affective and informational interfaces with which players interact have been distributed in a variety of forms, with contemporary computer, game console, and networked games representing only the most recent in a history of industrial game production that Collins (2008b) traced back to at least nineteenth century arcade games. The acoustic and musical capabilities of the game system on which a particular game is delivered has a significant influence on the quality of sounds that can be produced by the game and on the compositional techniques that can be used in the development of the game's soundtrack. In this section, I discuss game sound rather generically with respect to technology, emphasizing the conceptual roles fulfilled by sound within game play.

First and foremost, computer games are produced as part of a flourishing industry that has developed ties with other sectors of the entertainment industries. In this context, a primary function for sound in games is increased sales of games and, through allied marketing strategies, music, films, and other entertainment goods (Kärjä 2008; Tessler 2008; Collins 2008b, 107–122). Licensing of popular music for inclusion in game soundtracks is used as a strategy to reduce cost and risk in game development and as a strategy to promote the popular music used (Collins 2008b, 108). Having well-known film actors perform narration for games has been adopted as a strategy in top tier games to create the widest possible appeal for the game (ibid.).

The most elaborate of contemporary games divide game play into a sequence of episodes, some of which are linear and allow for soundtrack composition techniques similar to those used in film, subject to user interruptions only in the possibility that the game or the section of the game may be stopped or turned off. In other episodes, game play is driven by interactions between a player and the game state and the soundtrack along with the visuals must respond appropriately to an unpredictable, non-linear narrative structure. Introductory sequences, for a game itself or for episodes within a game, may include visual titles and *cinematics*, film or animated sequences inserted to provide information for the player, either of which will typically, argued Marks (2001, 188–191), be accompanied by music intended to set a mood and help to keep a player motivated to continue with the game.

Collins (2008b, 127–133) argued that, although most of the functions of film sound are employed in games to a greater or lesser extent, there are both difficulties in applying some of those techniques to non-linear game play and there are opportunities to use sound to prepare a user for imminent events, to give them information on which to make decisions and act, and to motivate them to continue with the game. Consistent and similar uses of game sound within loosely categorized genres of games can assist players in learning new games more or less quickly, depending on adherence to learned player expectations (ibid., 124).

A user's role in deciding the unfolding of the game narrative in a particular instance of play emphasizes the informational and motivational functions of game sound. Diegetic sounds with no visible on-screen source can cue a user to opportunities or dangers within the game (ibid., 133). Sounds can draw attention to important objects within the narrative and thereby help a user to filter the information presented during game play (ibid., 133). Sounds and music can provide thematic connections between game episodes to emphasize continuity within the narrative and help to re-establish the narrative in a game with which a player may engage over extended durations before completion (ibid., 131-132). Music can provide motivation, especially in relation to user-driven elements of some games, establishing rhythms and helping a player to physically achieve a target rate of play (ibid., 130). Similar to film sound, game sound can mask distracting noises which takes on additional importance because of the role of user decision making and response in the advancement of narrative (ibid., 132). Finally, game sound and music can signal emotional cues such as peril or sadness as a form of feedback to the user concerning impending game states, dangers, or the results of just-completed actions (ibid., 133).

As in film, sounds heard during game play can be understood to be either diegetic or non-diegetic. As argued by Collins (2008b, 125–127), the non-linear, player-driven sound design of games complicates the sound space and she differentiated game sounds along dimensions of *dynamics*, *interactivity*, and *adaptability* in addition to locating them within or outside of the narrative space. Dynamic sounds respond to either player actions or changing game states in some way rather than being repeated unchanged no matter how often played, as is the case with the score of a film or the soundtrack accompanying fixed game cinematics. Dynamic sounds, those that change in response to game play or game state, can then be differentiated according to how able a player is to directly control those sounds. Sounds that change in response to player actions are interactive and sounds that change in response to game state changes such as the passage of time within the game world are called adaptive. Games can be designed to use sounds that combine adaptability and interactivity or may vary sounds only in relation to player actions or game state changes. The characterization of sounds in a game design according to any of these dimensions may be fluid throughout game play with sounds that were static becoming dynamic at certain times, moving in and out of the diegesis, and possibly being interactive and/or adaptive only in certain situations or at certain times during game play.

The non-linear and indeterminate structure of game play in relation to time creates a novel compositional setting for music and sound design because the exact temporal trajectory of game play is unknown when a game's visual and soundtrack designs are being finalized. This indeterminacy requires a flexible soundtrack design and the ability for a game's software to initiate the playback of soundtrack elements and to control the playback of those elements according to parameters derived from the current state of an instance of game play. Guerraz and Lemordant (2008, 62) conceptualized the game soundtrack as being created from a set of *media chunks*, a contiguous region of audio data treated as a minimally playable sound unit, each of which is organized as part of one or more *cues*. A cue is a predefined event including the initiation of acoustic or visual phenomena within the game interface but also possibly including more abstract actions such as setting variables to influence future game states. Cues are triggered by *cue requests*, each of which is an indication that a particular game state has been reached.

Guerraz and Lemordant's description of game control structures for sound rely on all of the characteristics of digital media, as proposed by Manovich (2001) and discussed earlier. In particular, this way of thinking about game sound highlights the importance of modularity, automation, and variability as characteristics of the audio data that allow it to be manipulated, combined through dynamic layering, and repeated as necessary in response to game algorithms to support the establishment of narrative and in communicating elements of that narrative to a player, and thereby encouraging that player to act and further propel the narrative. Cue requests form bridges between the game's algorithmic model and events and the visual and acoustic interface elements through which a user comes to understand game objectives and the status of their current game play.

Because the duration of a particular instance of game play is also indeterminate and may range from very short to extended, probably intermittent, periods, the soundtrack and musical compositions for a game will most likely include elements that repeat. Reuse of soundtrack elements and music is common for compositional reasons and for reasons of cost in storage of audio data. Repetition occurs so commonly as to cause many people to detest game audio, especially non-players who happen to be close enough to overhear the game sounds without benefitting from the informational and affective benefits they provide to players. In recognition of the difficulty of avoiding repetition of media chunks in the design of game soundtracks and in an attempt to avoid tiresome game soundtracks, Kaae (2008, 84) argued that both game sound *adaptability*, similar to Collins's use of the term, and game sound *variability*, entailing the use of compositional and performance techniques enabling melodic and harmonic variation of game music over time, should be studied. Paul (2008) discussed the use of granular compositional and performance techniques as a group of methods that increase the variation of soundtrack elements during game play by allowing for the generation of sound effects and music during game play based on prior production of suitable, very brief sound fragments called grains.

Audiovisual cartography can draw much from research concerning game sound while recognizing that there are differences between the two, most notably the centrality of the player with respect to game narrative in the latter and the less certain role of narrative as a factor in the former, depending on subject matter. Designers of sound for games have begun to address problems concerning the management of sound cues within non-linear, indeterminate narrative structures (Guerraz and Lemordant 2008; Kaae 2008) but it is not clear that audiovisual map applications will often be constructed using similar structures to those of goal-oriented narratives that shape player motivation within a specific mission or that then gauge player performance against that mission.

By contrast, research concerning the management of audio as digital media within non-linear, indeterminate game narrative is directly applicable to interactive audiovisual cartography's requirement to been seen and heard to react to user actions. Interactive audiovisual cartography similarly must be able to vary both the visual display of the map and the acoustic accompaniment as dictated by the data model, for which the map acts as interface, and the user's actions selecting the aspects of the data model that should currently be displayed (Brauen 2011). Adaptation will be required depending on the authoring tools and, especially, the technologies through which a map application is distributed and used. In particular, for this thesis, designers of maps distributed and used as World Wide Web applications would not yet be able to assume that all computers upon which the map is used are capable of executing recently proposed and developed algorithms, such as granular synthesis or highly accurate beat matching when layering audio chunks containing music (either because of hardware or software differences between computers).

4.3 Human-Computer Interaction: Auditory Representation

Studies by HCI researchers of *sonification*, the algorithmic transformation of data into acoustic sequences, have been conducted to determine whether or not this can play a role as part of effective interfaces supporting process monitoring, pattern detection, and comparative analysis of trends within data.[8] Sonification has been used both as a vision replacement strategy for the visually impaired and as an additional modality with visual displays in interfaces where the visual 'channel' is already extensively used, such as in real-time navigation displays. Kramer et al. (1997) argued that the research agenda for sonification should encompass the following concerns: psychological and perceptual research concerning sound; development of software and utilities for the creation of sonification systems and applications; and the development of sonification applications to test and develop the other two areas.

A range of types of sounds have been studied for a variety of computer interface applications. *Auditory icons* (Gaver 1994), designed sounds mimicking sounds from the world, have been associated with cues within a computer interface to alert users to system events (e.g., the sound of crumpling paper to designate a file being put in the trash). *Earcons* are synthesized tone sequences intended to communicate information to users about computer interface objects and events as learned abstractions (Blattner et al. 1989; Brewster et al. 1994). Whereas auditory icons are argued to be more intuitive because of their similarity to real world sounds, Brewster et al. (1994) argued that earcons are combinable (through concatenation or hierarchical property inheritance) and, therefore, potentially more efficient.

Flowers and Hauer (1995) compared the abilities of participants interpreting time-series data presented using synthesized tone sequences to those of participants working with visual representations and found that the two groups performed about equally when asked to determine data characteristics such as slope

[8]Kramer (1994c, 186) differentiates *sonification*, and *audification* which he defines as involving a more direct transformation of the data into the audible domain, for example, by directly scaling data values to be used as frequency values in the audible range. Sonification, as defined by Kramer, involves the intermediate use of a sound generator with control parameters such as amplitude or pitch adjusted according to the data values.

or slope change. Flowers et al. (2001) proposed the use of multivariate acoustic representations of weather records for month-long periods, argued that the approach was well suited to the task of summarizing temporal data sets, and discussed some of the design choices made including mimetic use of melodic patterns to represent weather events and the transposition of temperature highs and lows to high and low musical pitches.

Gaver et al. (1991) studied the ability of test participants to collaboratively monitor the machines and processes in a simulated bottling plant, with each pair of participants running the plant for two one-hour sessions, once with a combination of visual displays and mimetic sound effects and once just using the visuals alone. The need for collaboration was enforced in each group by the test design that partitioned the process such that no participant was able to see the entire process but in the sound-enhanced case the entire process was audible. With the additional information provided in the audiovisual case, and a sound design that effectively signalled most of the process problems, they found sound to be useful, allowing participants to acoustically detect process problems and alert their collaborators even when their visual display did not include the process step that was failing.

Sonification has been used to convey information for visual impaired people. Theofanos and Redish (2003) developed a screen reader to narrate the contents of web sites for visually impaired users and screen reading software is now a standard feature in some home computers. To assist visually impaired people navigate their environment, Meijer (1992) developed a visual prosthesis combining digital capture of an image stream from a hat-mounted camera and real-time transformation of those images into tone sequences using computer-based signal processing to provide an alternative mode of sensory information for the wearer. Some users reported that after using the system for extended periods of time they began to think of the process in terms of 'seeing.'

Flowers (2005) argued that despite advances in desktop computing allowing for complex manipulations of multiple acoustic streams, research into sonification still often tended to overestimate the capabilities of human auditory perception, attention, and memory, especially with respect to handling multivariate representation. He further argued that sonification research often tends to underemphasize "that, like visualization, optimal display formats may be highly dependent on the specific task context to which they are to be applied" (ibid., 406).

Research into cognitive processes underlying perception and thought indicate the need to consider not only the perception and understanding of each sensory modality but also the interactions between modalities. Despite the importance of human spatial and visual capabilities, debates continue concerning how our minds represent and manipulate spatial and visual information. Debates have tended to cast mental imagery as being represented either in a picture-like manner or in a language-like manner (see Block 1981; Tye 1991). Paivio (1986) argued that human mental processing separated imagery and coded information, such as language, through the use of separate processing channels, each having capacity limits. Reisberg and Heuer (2005, 45–48,52–53) argued that evidence exists that mental imagery functions as both depictive, representing the imaged form pictorially, and descriptive, containing spatial information concerning the organization of the imaged form. Mental images are thus considered to be functionally similar to external pictorial representations in their ability to depict, but that depiction is augmented with additional descriptive information including structural and spatial relations. In addition, Reisberg and Heuer (2005, 44-45) argued that the processing of mental imagery and visual perception probably utilize the same parts of the brain based on empirical evidence confirming that mental imagery can disrupt visual perception (and vice versa).

Use of pictorial information is not isolated from other sensory modes. Pictorial information and coded forms of information are often used together (maps being a good example). Mayer and Anderson (1991, 1992) argued, based on their experiments with learners using educational media, that effective learning outcomes require that perceived verbal information (whether read or heard) be incorporated along with visual information into a representation containing verbal representations, mental imagery, and referential linkages between the two. The possibilities of interference between visual perception and mental imagery, theories concerning the referential relationships between verbal representations and mental imagery, and the understanding of both the visual and auditory perception channels as having limited capacity (Baddeley 1986) led Mayer (2005) to propose a number of guidelines for the design of educational multimedia. These included a suggestion to offload a learner's vision by providing narrated verbal information rather than visual text accompanying visual media and a suggestion to use synchronized narration along with visuals rather than sequential narration preceding or following the visuals.

In developing research agendas for audiovisual cartography, HCI studies may usefully highlight guidelines concerning the communicative effectiveness of sound and sound properties when intended to enrich users' abilities to understand complex data. As Flowers (2005) argued, some lessons have been learned concerning how sound can and can't be applied to effectively enable certain tasks. But some of these studies also focus on using sound and its properties according to a fairly narrow perspective, largely ignoring the cultural attachments possible with sound, possible affective responses to music, voices, and sound effects (as opposed to 'sound' as a generic perceptual mode), and the connotative suggestions possible with this range of types of sounds. Although Flowers et al. (2005) argued that the acoustic processing power of desktop computers and experience with sonification enable expert users of "specialized applications" and "are easily comprehended by normal users with minimal practice," it is important to note that the method of listening and thinking about data suggested by these studies are not methods with which untrained users can be assumed to be familiar. Training, guidance, and assistance may be required to help users understand the intention of these designs.

This is not to argue that sonification is incompatible with other perspectives on sound design. Rather sonification and sound design for film and games are all perspectives that can provide insights into the the use of sound for audiovisual cartography. Research is needed concerning possibilities for assisting users to understand and work with audiovisual maps, encompassing training, self-guided help features, and design guidelines for the maps themselves.

—

The interaction between sound and space in each of the disciplines addressed in this chapter is not homogeneous. Film has tried various approaches to reproducing diegetic sounds, some intended to match the acoustic and visual point of perception and others not. Certain game genres have developed styles of positioning the player with respect to a game world (e.g., first-person viewpoint of racing games, overview viewpoint of simulation and strategy games) but the use of spatially processed sounds is still rare, partly due to the cost of processing audio to suggest three-dimensional positioning and partly due to a tendency within the industry to continue to focus on visual design (Hamidi and Kapralos 2009). Within the purview of HCI research, studies concerning spatial perception of sound (Shinn-Cunningham 2000) and of implementation and design issues for the creation of software and systems supporting the spatialization of sound have been under-taken (Cohen and Wenzel 1995; Wenzel 1994). How spatialized sound would be used in different applications is beyond the scope of this discussion. Relationships between visual and acoustic space as interpreted by users of audiovisual map applications is an area requiring future research.

5

Guidance from Dynamic Cartography for Interactive Audiovisual Cartography

This chapter examines the objectives and methods of animated and interactive cartography, referring to both as *dynamic*[1] cartography. This approach argues that similar types of geographic analyses may be amenable to the explicit use of the temporal dimension whether represented visually, acoustically, or through a combination of sensory modes. The objective of this discussion is to understand how dynamic cartography could inform possibilities for audiovisual cartographic design. Section 5.1 briefly reviews literature concerning animated and interactive cartography, highlighting the motivations behind research into these topics. This review then forms the basis, in Section 5.2, of a detailed decomposition of dynamic cartographic artefacts into a set of abstract functions and an appraisal of these functions to determine which if any are amenable to representation using sound.

Why Examine Dynamic Cartography for Insight into Audiovisual Cartography?

Animation and interactive computer applications share an explicit temporal di-

[1] I will refer to interactive and animated cartography as 'dynamic' cartography because of the non-static nature of the images (and sounds when audiovisual cartography is also implied). I will use the terms 'interactive' or 'animated' as needed to distinguish between them.

mension with sound. But is this enough to assume that guidance for audiovisual cartography can be derived from predominantly silent media such as strictly visual maps? I think there are several reasons why such an examination is warranted: the potential to apply audiovisual cartography to problems that are similar to those already studied by research into dynamic cartography; audiovisual cartography, beyond instructional narration intended as an overview for a map, is necessarily interactive or animated; and research into the effectiveness of audiovisual media has suggested that care be taken in understanding the linkages between sensory modalities as well as the use of each independent modality.

First, as I will discuss in more detail in Section 5.1, one of the reasons given for researching the use of visual dynamics within cartography has been the desire to use representational dynamics to examine and present situational dynamics abstracted from complexities of reality. Thus if there is potential for sound to be used as another sensory modality in cartographic representations then the problems to be examined may be many of the same process oriented problems for which dynamic cartography has been studied. Similar pitfalls and strengths may or may not emerge but an examination is warranted.

Second, although sound envelops us in three-dimensional space as has been discussed in relation to film (see Chapter 4), it is also a time-based medium. Sound requires time for waves to form, travel, and be perceived and interpreted. To use sound as a representational modality within maps, as part of the construction of digital codes signifying discontinuity and difference within an analogical coding of space (Casti 2000), requires a means to select the criteria for representation, including location (or region), time period, and thematic criteria.[2] The digital auditory

[2] There are many possible means by which the map area or location associated with the current selections and settings of the sound design could be selected. For example, the cursor may designate a point location, linear feature, or areal feature. Alternatively, pan and zoom controls for the

codes then signify the mapped differences over time, and in relation to other possible selection criteria, as a sequence of criteria are selected and the acoustic changes caused by those selections are perceived. In each design, ensuring that a user can interpret the selection criteria to which a sound design currently relates and the means by which a user selects a sequence of such criteria over time to be represented are important considerations. Thus the use of sound to represent spatial data within maps necessarily requires some sort of criteria selection (locational, temporal, thematic) to occur and this is most easily conceived through either interactive selection or configuration of animated sequences.[3]

Third, learning and deriving information from an audiovisual map, which as has just been argued, will most usually also be an interactive or animated audiovisual map, will require a user to correlate sensory perceptions spanning vision and listening (while also, at least some of the time, coordinating gestures to initiate criteria selections via an application interface), as was discussed in Section 4.3. Audiovisual maps may benefit users by offloading some information processing to a user's auditory processing systems, but cognitive load across sensory modalities is an important issue requiring study.[4] An examination of sound uses within cartography should attempt to understand the integrated use of audio and vision.

---

visual map may select the region over which spatial variables are computed, resulting in updated audio selections and settings.

[3]This discussion does not rule out the use of introductory narrations relating information relevant to the entire map or the use of the map application but such use is not the focus of this discussion. Expository narrations and sound design for audiovisual maps are probably complementary and it should be possible to combine their use within a single application.

[4]The literature on cognitive load focuses on the issue with respect to *learning* and there are reasons for employing sound other than the impartation of information. For example, sound designs that are predominantly intended to increase user engagement without directly increasing information content. Although the cognitive load literature may not directly address such usage, the paucity of sound use in maps may mean that users spend effort trying to ascribe meaning to the sounds, at least during important initial use when they are trying to understand the overall map design. Film literature concerning interactions between sound and imagery in shaping meaning (Chion 1994) similarly suggest that segmenting sound use into *information imparting* and *non-imparting* cases is problematic.

5.1 Research Motivations for Dynamic Cartography

Although sound has not been ignored in research concerning the applications of dynamic media to cartography, as outlined in Section 1.3, the focus has unsurprisingly been on visual dynamics. Animation and interactive visual cartography have been the predominant focus, although tactile and audio-tactile maps have also received attention (Golledge 1991; Jacobson 1994; Rice et al. 2005; Siekierska et al. 2003). Focusing on the visual, there has been considerable research concerning both animated (Dorling 1992; Monmonier 1992; Peterson 1995; Thrower 1959), and interactive (Crampton 2002; Monmonier 1989) cartography as means to provide alternative representations to those possible with a static map.

Sound has occasionally been mentioned as a promising addition to the techniques of dynamic cartography (Monmonier 1992; Peterson 1996; Thrower 1959) but a systematic focus on the use of sound appears to have been rare (cf. Krygier 1994). Furthermore, there appears to have been no examination of cartographic sound usage that asked whether or not some of the models and metaphors that have been adopted by researchers working on the visual aspects of animated and interactive cartography could be applicable to sound. What has been the motivation for the representational models developed in research concerning animated and interactive cartography? What metaphors and models have emerged from this research and which, if any, of these are applicable to the use of sound in audiovisual cartography?

5.1.1 Animated Cartography

Animated cartography has been a topic of research since at least the 1950s (Campbell and Egbert 1990; Thrower 1959). One appeal of animation for cartographers is the ability to represent spatiotemporal processes using temporally dynamic media such as film or computer displays. Although techniques such as arrays of multiple small maps or the use of isarithms have been developed to represent processes, especially temporal change, in static cartography, the animated map presents a sequence of images at a sufficient rate for the represented phenomena to be perceived as a spatial process rather than as a sequence of snapshots. According to Peterson (1995, 6), the use of animation draws on humans' abilities to detect visual change and movement which are possibly more developed than our abilities to recognize static images. "In a sense, what happens between each frame is more important than what exists on each frame" (ibid., 48).

Models of cartographic animation have been developed from the perspective of how each successive frame is created and then assembled into an animation (Peterson 1995) and from the perspective of technologies that can be used to create the animated sequence (Gersmehl 1990). The latter classification provides greater detail but the underlying models are the same as those outlined by Peterson (1995, 143–152): *frame-based* or *cast-based.*

Frame-based animation uses some means of creating a set of final images for each frame of an animation which are then assembled into a sequence, possibly in conjunction with audio. Early cartoon motion picture animation was framed-based. A series of maps showing the expansion of a city's urban boundary at one-month intervals assembled into an animated sequence would be a cartographic example of frame-based animation.

Cast-based animation uses a cast of visual objects which move independently over a stationary background (or a series of stationary backgrounds representing different scenes or frames). Each cast object is called a *cel* because the original techniques of cast-based animation implemented each independent object as a sequence of images on celluloid that were pegged onto the background image for photographing. The complete cast was assembled on the background for each frame using layers of celluloid. Although the original photographic technique did create a series of final images that were sequenced, contemporary computer-based animation software, such as *Adobe Flash* or the World Wide Web Consortium's (W3C) graphics specification language *Scalable Vector Graphics*, maintain the independent object metaphor both in the authoring tools and in the display technologies. Television weather maps are a common example of cast-based cartographic animation in which weather phenomena such as wind, rain, and cold fronts are implemented as cels over a stationary map image. Monmonier (1992) discussed the use of graphic scripts to present thematic spatial data in a manner that treated the map, associated bar charts, and statistical plots as synchronized cels.

Although animation is most commonly employed in cartography to demonstrate process change over time, other orderings of animation have also been proposed. David DiBiase et al. (1992) used the term *re-expression* to denote the presentation of an animated sequence in an alternative ordering or using an alternative pacing of the animation other than presentation time as a linear scaling of phenomenal time. The ordering could be based on values of a represented spatial variable (e.g., hurricanes ordered by severity of annual impact) or could use an altered pacing of the sequence in which the duration of each section of the animation is derived from a variable (e.g., the total number of hurricanes within a region during the time depicted). Taylor (1987) sequenced a display of literacy levels for India's states (as classified percentages of the population) so that the outline of India was thematically filled in as groups of states assigned to descending classification levels were coloured on the map. Re-expression may provide additional opportunities for a viewer to determine whether or not spatial patterns exist in data.

Monmonier (1992) proposed techniques that quickly change a cartographic visual display to draw a user's attention. *Blinking*, in which one or more symbols are flashed for a period of time, he found useful in drawing a user's attention to specific items within a display. Another technique, subsequently referred to as *flickering*, alternately displays each of a pair of spatial variables mapped over the same area, allowing a viewer to look for patterns of correlation between them. "Because quasi-movement reflects dissimilarity, the overall degree of visual dissonance is inversely proportional to correlation" (ibid., 252). Peterson (1996) extended flickering to more than two variables by sequencing city-level maps showing a set of related variables by district (percentages of population by age group).

Although re-expression and flickering are usually used to show process change over a fixed spatial extent, other applications of animation vary the spatial extent alone or in combination with changing the thematic content of the map. *Fly-through* animations (Moellering 1980) alter the viewpoint, often in conjunction with a simulated three-dimensional surface to emulate the experience of flying over a terrain.

Peterson (1996) proposed animations using sequences in which the mapped data remained unchanged but the cartographic representations changed to highlight decision processes in cartographic abstraction and to offer additional opportunities for a viewer to detect patterns in the data. Although some of these animations could be used with a fixed spatial extent (e.g., a sequence of maps showing the effects of different data classification schemes in choropleth mapping), others also include changes in spatial extent. As an example of this latter category, Peterson used the term *cartographic zoom* to refer to animations in which selection and symbolization change as the scale of a displayed map increases or decreases.

As is commonly done with static cartography, the selection criteria change along with scale allowing more types of information to be included as scale increases. Similarly, map features that can only be shown as points at small scale may be transformed into areas as the viewpoint zooms in. In recent years, this technique has become almost ubiquitous through applications of *Google Earth* such as visual backdrops for television weather and news reports or computer-based geographic mash-ups.

Kraak and Klomp (1995), summarizing Dransch (1995), described additional variations on the idea of animating a sequence of changing cartographic representations such as the use of different symbolization types (e.g., isoline, dot density, and choropleth). In addition, they discuss animations using *successive build-up* in which the sequence of images is designed to show a fixed spatial extent while separate thematic layers are added to the display in an attempt to help a viewer understand spatial context.

5.1.2 Interactive Cartography

Monmonier (1992) stated that a common request from people who had seen his graphic scripts prototype was for the ability to pace the presentation, reverse the direction of animated sweeps, and review introductory sections. Control of computer-based cartographic animations was then one motivation for the addition of mechanisms allowing a user to decide when, and how often, they viewed portions of a cartographic representation. But other factors such as the relative expense of producing paper maps at a variety of scales to examine particular spatial phenomena compared to viewing maps on dynamic computer displays also spurred research into interactive controls. Dorling (1992, 217) argued that interesting spatial patterns often exist at very large scales and "rapidly blur away as the spatial scale of analysis is reduced", requiring interactive tools, such as *pan* and *zoom*, to allow adjustment of the extent of the mapped area. The declining cost and broad availability of computer technology in the 1980s also prompted research into the use of computers for analysis of large data sets, spatial and otherwise (Crampton 2002; Dorling 1992). Interactive techniques developed for statistical analysis were adapted for spatial analysis, *geographical brushing* (Monmonier 1989) being one such adaptation.

Monmonier (1990) examined approaches to analysis of time-series data ranging from static maps and charts to use of interactive computer displays. He argued that, compared to static maps and graphs, interactive mapping and charting using dynamic computer displays offers a promising alternative for dealing with the complexity that arises in data sets comprised of multiple variables observed periodically at a number of locations (also see the discussion, in Section 4.1, of Monmonier's (1994) argument that *navigation* and *narration* are the two guiding metaphors for animated and interactive cartography).

Asche and Herrmann (1994, 217) examined the possibilities of utilizing digital technologies not only for map production and design but also for publication and use. The dynamic use of the map as an interactive interface, they argued, would support the "integration of time-based data, of animation, and sound" and the inclusion of interactive controls would allow both a cartographer and a user to experiment with alternative cartographic representations.

Krygier et al. (1997), Crampton (2002), and Persson et al. (2006) defined typologies of interactivity for the presentation of geographic information, each with different purposes and details. Krygier et al. (1997) proposed a typology of multimedia element types for use in the development and delivery of post-secondary courses in large lecture halls using computers and overhead projectors. Their typology categorized multimedia content to be used as part of a lecture according to *function*, describing how the instructor intended to present and pace the delivery of information for which the multimedia content was an aid, and *form* detailing a range of predominantly visual presentation materials. This is similar in intent to multimedia typologies that have been proposed by Heller and Martin (1995, 1999) for training multimedia designers by encouraging them to consider alternative types of media and different abstractions for working with media in presenting ideas. The range of forms considered by Krygier et al. is both more detailed in explicit consideration of alternative graphic forms, such as maps, and more restricted because it ignores explicit consideration of text and sound within the taxonomy except when included as part of an encompassing form.

Crampton (2002) proposed a typology of interactive user functions, intended as a basis for comparing the level of interactivity supported by geographic visualization systems, and included the following functional categories: *interactions with the data representation*, *interactions with the data*, *contextualizing interactions*, and *interactions with the temporal dimension*. While interacting with the data representation, a user manipulates the system to produce different views of underlying data. This could entail, for example, changing the cartographic representation used in maps, selecting additional feature types (often implemented as visual layers) for display, listing all or a selection of attributes associated with some or all mapped features, or adjusting displayed map extents.

When interacting with the data, a user queries and filters information from the system database to focus the presentation of information for the geographic investigation being conducted. Techniques such as geographical brushing and highlighting may be used to draw attention to the results of the queries and to assist a user in determining the next steps in the investigation.

In discussing contextualizing Interactions, Crampton (2002, 88), drawing on Monmonier (1992, 1994), argued that when using a digital mapping system, a "user's decision-making and choices are conditioned by what the user has already learned and what they still [...] would like to know." Despite the explicit temporal dimension in his discussion, he included in this category only interactions that provide multiple simultaneous lenses on the data being examined.

While interacting with the temporal dimension, a user interacts with dynamically changing (animated) maps or other graphic displays, with the possibility of controlling the sequencing of the displayed data representation (e.g., ordered by time period, or through the use of flickering or re-expression, as described in Section 5.1.1). Crampton also grouped fly-throughs and navigation in this category.

Persson et al. (2006) proposed an alternative typology of interactive functions found in geographic visualization applications, comprised of eight groupings, intended to encourage a standardization of map visualization capabilities and to enable comparison of existing visualization systems. Creating eight groupings, as opposed to Crampton's four, Persson et al. (2006) produced a typology that lists a greater number of functions and provides a finer grained classification among them (e.g., a separate category of functions for interacting with simulated 3D space).

For this discussion, key features of each of these typologies include: a focus on the ability of users to query data and focus geographic displays on subsets of information relevant to a current information need; a focus on the ability of users to link multiple representations of data together, building on earlier techniques such as geographic brushing; and a focus on the temporal aspects of a user's interactions with the application, either because of the processual unfolding of a user's interactions with a system or because of the inherently temporal characteristic of certain display techniques such as animation and flickering. Section 5.2 builds on these themes, explicitly using Crampton's typology to structure a discussion of possible methods for using auditory representation as a component of interactive audiovisual map applications.

5.2 Linking Sound to Mapped Space and Time

This section outlines an analysis of possible uses of audio for *linked* representations (see Section 3.1) of data along with visual cartographic displays, using an analogy to the linked visual displays of interactive cartography. Like linked visual displays, linked auditory representations are used to provide information or perspectives concerning geographic features or regions as a user interacts with the application. The analysis is structured using functional categories similar to those proposed by Crampton (2002), discussed above, and includes many of the interactive functions outlined by Crampton and by Persson et al. (2006).

The audio playback and processing model supporting linked representation throughout this discussion is that of a multi-track sound mixer and software synthesizer allowing simultaneous playback of multiple audio recordings, each contained within a *track*, along with playback of one synthesizer composition encoded using Musical Instrument Digital Interface (MIDI) Stored MIDI File (SMF) format (Heckroth 1996). Furthermore, the audio model allows independent control for each audio recording, and of playback and effects processing parameters associated with each recording (e.g., start, stop, looping, gain, left-right panning, mute), and for the synthesizer composition (e.g., start, stop, looping, gain, left-right panning, tempo).

Figure 5.1 provides a simplified depiction of the architecture of an audiovisual map created using such a model. As shown, the sound design for an audiovisual map may be composed of multiple recorded audio tracks (three are shown) along with one synthesizer composition, each of which may be independently played as indicated by the arrowheads marking current playback position for each. The synthesizer composition is constructed using multiple synchronized instrument parts, each of which can be independently controlled by adjusting acoustic parameters (e.g., gain, pan, mute, or pitch transposition).[5] To ensure synchronization between instruments, tempo parameters along with controls to start or stop the music are applied only at the level of the composition itself.

To be used as part of an interactive audiovisual map application, the audio playback and processing model provides an Application Programming Interface (API) that allows cue requests triggered by user actions (e.g., cursor movements on the map, clicking on hyperlinks, adjusting audio configuration controls) or application processing (e.g., progression of a sequential display of mapped data) to update the playback states of each track or the synthesizer composition (e.g., *stopped, started, unloaded*). Similarly, cue requests could update the configuration of acoustic parameters associated with each track or instrument in the composition. Thus the audiovisual map application can be designed such that acoustic updates, coordinated with visual application updates, can alter playback controls of some or all

[5]MIDI, prior to defining the SMF format, was strictly a messaging protocol that defined messages used to describe the performance of the parts comprising a musical composition (e.g., 'note C1 on with velocity 65', 'note F1 off'). The SMF adds timestamps that allow all of the notes of a performance part to be specified in advance as a queue of timed messages. The original default for SMF format was a *Type 0* file that included all channels within a single track and would not allow the separated, independent control of each instrument assumed by the audio model described here. Independent control of instruments within a MIDI SMF composition requires that the file is a *Type 1* SMF that stores multiple *tracks* of composition data, each of which can then be used to isolate a MIDI *channel* containing all of the instructions required to play a musical part including settings that select the type of instrument on which the part is to be played.

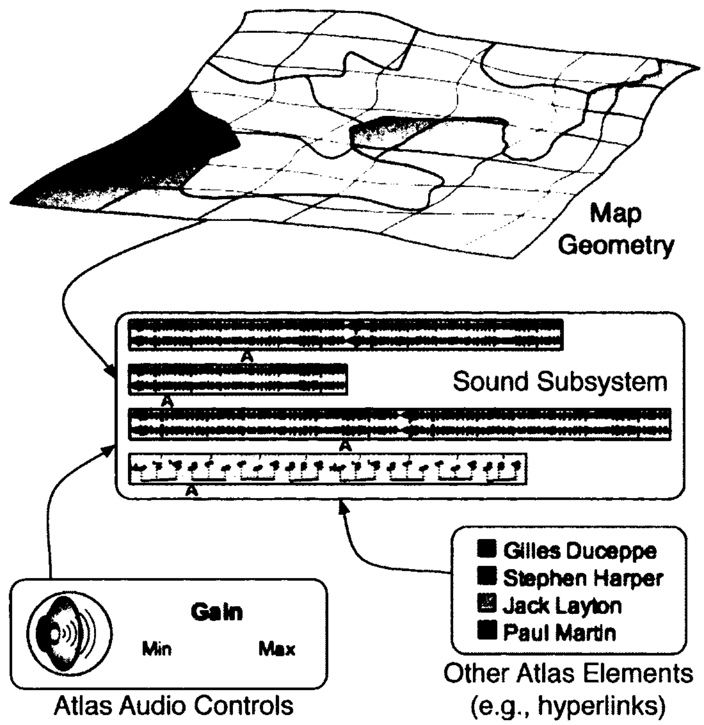


Figure 5.1: High level interactive audiovisual map architecture (figure concept and design: Glenn Brauen; the map sheet image is adapted from an image posted on http://www.openclipart.org by Ted Balmer and placed in the public domain).

sounds in use by the application. Each sampled audio track and the synthesizer composition may be played either in looping or non-looping modes, determining whether or not the sound restarts from the beginning once it has played to the end.

A software-implemented sound subsystem, consistent with this model and appropriate for use as part of audiovisual maps distributed and used over the World Wide Web, is described in Chapter 7 and Appendix F.

5.2.1 Guiding Metaphors

In analyzing interactive and animated cartographic functions for applicability to audiovisual cartography, the following metaphors are proposed for the design of graphical and linked acoustic representations of spatial data. Case-specific design and testing are required to ensure that a user can determine whether or not a specific metaphor applies when interpreting an application's audiovisual representation.

Sounds are caused by events: Sounds are always created by the generation of a pressure wave in a medium suitable for carrying that wave (e.g., air, water) with a frequency in the range audible to a listener. Sound is always the result of an event. Learned responses to sounds such as being startled or turning to look in the direction from which a sound came are a result of the understanding that it was caused by an event. Therefore, within audiovisual cartography a certain set of sounds can be reserved for use as alerts that something has happened or to draw a user's attention to a particular element of the visual or aural representation. Principles of *auditory grouping* (Williams 1994) are used in interpreting sounds as being caused by new events or part of existing event sound streams.[6]

Representational change indicates phenomenal change: Extending the above event metaphor, sound causing events of differing types or similar sound causing events of differing magnitudes often cause distinct sounds. Therefore, acoustic changes occurring as a result of quantitative or qualitative variable changes can be a simile for the behaviour of sounds in the environment: sounds change as a result of events happening differently. Ballas (1994, 84–85) argued that "[s]imile and metaphor are the core of sonification" and that certain visual/aural relationships were well understood by users of auditory representations, although variations according to listener characteristics such as age and level of musical training exist: "sound frequency with vertical placement, wave form with pattern, amplitude

[6]Williams (1994) argued that Gestalt principles apply to sound grouping and identification. While identifying separate sound events and stream continuations, listeners assess characteristics of sounds and variations of sounds over time, with Gestalt principles being applied, such as: similarity, proximity (e.g., in frequency and loudness), good continuation, and common fate (Levitin 2006, 73–80; Goldstein 1999, 359–364).

with size, and duration with horizontal length." The parameter updates causing acoustic and visual changes must be understood by the user and there is potential for variation in the application of this guideline. For example, when using a continuous sound, it may be desirable to introduce some level of acoustic variation not directly related to the changes in the underlying data to avoid having the sound become tiresome. Such variations to avoid tedium should ideally not interfere with the interpretation of data-driven variations in the sound.

With respect to mapping data values to acoustic parameters, a partial, tentative list of parameter mappings, not intended to be prescriptive, include: 'louder means more,' 'higher pitch and tempo means happier,' and 'faster means more.'

Louder means more: increasing the loudness of a sound seems to be an appropriate response to an increase in the value of a quantitative variable. But Flowers (2005) cautioned against the use of loudness for continuous sounds because of habituation, the perception that a steady-state sound is not as loud or noticeable after a period of listening, and the relative perception of loudness when a sound is combined with others.

Higher pitch and tempo means happier: although there are many interacting characteristics of a piece of music that could influence user interpretation, generally a piece of music with a quicker tempo that remains in the mid- to upper-registers will be understood to transmit a lighter or happier mood than a slower melody that ventures into the bass register. For some maps, happiness, possibly in relation to wellness, could be a direct result of modelled effects. In other cases, this metaphor may assist users in remembering qualitative associations of the mapped phenomena (e.g., warmer temperatures) depending on audience. It is possible that the use of the higher register or the quicker tempo would each alone signal the happier mood. Other characteristics of a melody such as mode and the timbre of the instrument the melody is played on could also affect interpretation.

Faster means more: A sound with a higher tempo than a similar but slower sound contains more acoustic events per unit time and therefore seems appropriate to indicate an increase in the value of a quantitative variable, especially if it can be interpreted as a rate.

Some of Ballas's suggested relationships are consistent with the metaphors proposed above but the audio playback and processing model proposed for this discussion would in some cases only be able to approximate these relations. For example, the MIDI synthesis method most easily provides step-wise frequency increments because, ignoring the use of pitch bend, it emulates a keyboard with defined frequency relations between notes. These may be appropriate for a classified mapping between the data and acoustic domains but would only be an approximation of the frequency/vertical placement simile suggested by Ballas if used without classification.

5.2.2 Functional Analyses of Linked Auditory Representation

Table 5.1 provides a summary of functions drawn from the analysis of animated and interactive cartography discussed in Section 5.1, grouped according to functional categories derived from Crampton (2002). For each function, the table indicates whether or not there is an analogy that could be or has been applied to the use of linked sound in audiovisual cartography. Brief explanations of each function follow, tagged with the function reference from Table 5.1. As is discussed briefly for some of these functions, there are cases in which the only way I can imagine sound being used in a manner that would be analogous to the visual function would be by using either synthesized speech or narration to read out information that is essentially encoded textually. Although there are undoubtedly cases in which such an approach would be useful, for the purposes of this discussion focused as it is on the linkage between auditory and visual components of maps, these functions are encoded in Table 5.1 as 'not applicable' (✗, see table caption for the coding legend) as analogies and an explanation is provided in the discussion.

Again it should be stressed that Table 5.1 and the discussion supporting it represent a survey of *ideas*, some of which remain to be fully tested. Careful audiovisual design and tests of the resulting function within an application context may be required for each of those that have yet to be fully developed. The discussion attempts to distinguish between those that have and haven't yet been demonstrated.

5.2.2.1 Interactions with Data Representation

(R1: ≈ or ✓) As derived from film technique, the frame defining functions 'pan' and 'zoom' are inherently visual and have been translated for use with respect to interactive maps as analogous map framing functions. Although frame definition seems to be an intrinsically visual function, 'panning' has been adopted as an audio term describing an adjustment of left- and right-channel sound levels that apparently moves a sound left or right in an acoustic field with respect to a listener. Similarly, asymmetric sound pick-up characteristics of microphones along with delay differentials for sound pick-up by each in an array of multiple microphones has been used as the basis of 'acoustic zoom' technology for broadcast sound production and for remote microphone control to isolate the current speaker in conference auditoriums (de Lange 2010).

Although these acoustic uses of the terms 'pan' and 'zoom' can be thought of as analogous to their visual antecedents, there are different possibilities for the application of each of these visual concepts to linked audio representation. Acoustic

Table 5.1: Applicability of visual cartographic functions as metaphors for linked auditory representation: ‡ ≡ a fundamental component of this proposal (see text); ✓ ≡ applicable; ✗ ≡ not applicable; and ≈ ≡ similar concepts can be applied for acoustic representation but not to achieve the same goals.

Function	**Applicable?**
Interactions with Data Representation:	
R1) Map framing (pan and zoom) functions	≈ or ✓
R2) Map measurement functions	≈
R3) Adding or removing mapped feature types (layers)	✓
R4) Identifying meanings/legend function	‡
R5) Listing a set of parameters for a specified set of areas or features	✓
R6) Listing all or additional parameters for a specified area or feature	✗
R7) Comparing observed values to statistically significant values (e.g., maximum, minimum, mean, median)	✓
R8) Modifying symbolization or classification method	✓
R9) Emphasizing particular data through blinking	≈
Interactions with Data:	
D1) Knowledge Discovery from Databases (KDD): searching for patterns	✓
D2) Logical combination of data layers and generation of new variables	✗
D3) Examination of data and of results of interactions with the data: cartographic cross-classification (brushing)	✗
Contextualizing Interactions:	
C1) Linked displays	‡
C2) Many simultaneous views of data (small multiples)	✓
C3) Dynamic comparison: multiple representations of data	✗
C4) Dynamic comparison: representation of specified observations or of observations in specified orders (re-expression, and flickering)	✓
C5) Dynamic contextualization: successive build-up	≈
C6) 3D navigation or fly-throughs	✓
Interactions with Temporal Dimension:	
T1) Dynamic comparison: representation of temporally ordered observations	✓
T2) Manipulation of dynamic variables (e.g., duration, period)	✓
T3) Playback controls (e.g., stop, play, repeat) and manipulation of the time axis (stretching or compressing)	✓

pan does not perform the same function of selecting that which is included in a representation. Rather, it apparently moves that which is already in the representation to a different location relative to the listener. This can be used to separate sounds across the sound space created by stereo or surround speaker systems to ease auditory separation for a listener. Spatializing certain acoustic alerts, even with a simple left-right pan parameter,[7] could be a useful means to draw attention to certain areas on the visual map (this is also discussed below in relation to the visual function known as *blinking*, R9).

The acoustic usage of zoom as described above does parallel the requirement to decide which spatial extent or location in a map is to be represented by linked audio. The linked audio could be computed in some way over the currently visible map extent, in which case the visual pan and zoom functions do the linked audio selection as well, but this approach has not been used in any of the map examples discussed (or evaluated) in this thesis. Instead, the approach used by the audio-visual map applications discussed here has been to use the computer cursor as a 'virtual microphone boom,' designating a point, line, or region which is then represented using sound (and visually marked to help the user see the relationship). Extending the difference between visual and auditory space to the map metaphor, it would be possible to represent locations or areas beyond the visibly mapped region using linked auditory representation but the means by which a user is shown what the linked region is would need to be carefully designed, possibly through the use of a smaller scale reference map.

(R2: ≈) Map measurement functions do not have an obvious auditory analogy that achieves the same goal of returning numerical distances, although it would

[7]The sound subsystem to be discussed in this thesis (see Chapter 7) provides left-right panning controls but does not support three-dimensional spatialization of sounds.

certainly be possible to design a measurement function that narrated the results. In general, as discussed above, narrating fundamentally textual information would be possible for many of the functions listed in Table 5.1 but I am not going to repeat this for each. It would be possible to design a map interface in which sounds were parametrized based on computed distance(s) between the cursor and one or more point locations. For example, adjusting playback delays for each of two acoustic pulses to be proportional to the current distance between the cursor and each of two points, one associated with each of the tones, could be used to assist in locating a point half-way between the two locations.

(R3: ✓) Adding or removing feature types or layers in a visual map may be done to understand the context of the data displayed or as a means of searching for spatial relations between different variables in a data set. Although these operations may often be accompanied by other analyses to determine whether or not, for example, a statistical correlation exists between variables, the addition of visual layers is analogous to the creation of additional, coordinated auditory representations. During subsequent examination of a map, for example through the use of interactive re-expression, a user can then attend to the correspondence of the auditory representations of some or all of the included variables to determine if associated sounds tend to vary similarly as they work with the map. The suspected relationship between variables may be used to select representations for variables to be compared to enhance that comparison. For example, using rising pitch or scale transposition on different instrument voices (e.g., violin and vibraphone) may highlight the regions in which the associated variables increase in unison or emphasize the lack of spatial correlation.

Representing multiple variables simultaneously using linked audio creates design challenges with respect to how deeply you can 'stack the layers'. The ability of users to comprehend the data associated with mapped features needs to be considered. Both acoustic and visual representations become difficult when too many layers or variables are simultaneously mapped. Textual lists of feature-associated data allow an unlimited number of attributes to be presented but users cannot easily comprehend patterns in such lists. Sound design for film and games deal with concurrent cues organized within narrative structures and some human-computer interaction research has combined layered and sequenced sound cues to signify complex data using composed sounds, for example, concerning 24 hour weather forecasts (Hermann et al. 2003).

User interface controls to adjust the loudness of each linked auditory representation can be provided to allow a user to temporarily fade one or more of the sound layers up or down to allow them to focus on a particular set of variables, to adjust for peculiarities in their speaker system, or according to listening preferences. Silencing each variable is analogous to turning off visual layers.

(R4: ‡) A means of learning how a linked auditory representation works in coordination with the visual map design is a fundamentally important aspect of the audiovisual design and usage process proposed here. Figure 5.2 shows an example of a dynamic audiovisual legend design developed, as part of this research, for an audiovisual map application called *Canada-USA Commodity Trade 1976–2000*. The legend is designed to show a region's contribution to total Canada-USA trade as a percentage, and to be used a) as a direct audio controller allowing a user to select different values off the scale to hear how the sound design responds; and b) as a visual read-out coordinated with the sound design to assist users to understand or verify the meaning of the sound design as they work with the map. Dynamic audiovisual legends can be understood as an extension of research concerning *active legends* for cartographic interactivity (Peterson 1999) and *dynamic legends* to accom-

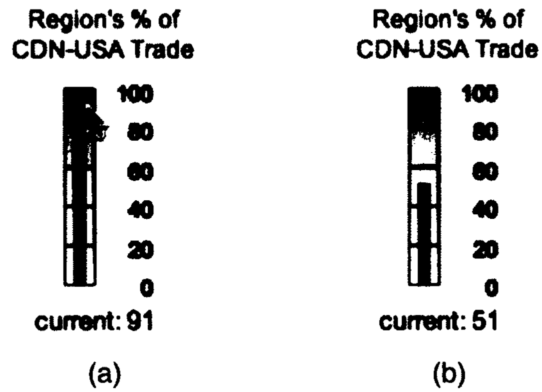


Figure 5.2: Dynamic audiovisual legend: a) as direct audio controller, and b) as a visual read-out and aid for learning the behaviour of a linked auditory representation (figure concept and design: Glenn Brauen).

pany temporal animated sequences (Buziek 2000; Kraak et al. 1997).

This legend and other examples of dynamic audiovisual legends are discussed more fully later in relation to the applications of which they are a part. One application incorporating dynamic audiovisual legends, *Canada's Trade with World Regions: 1976–2000*, was evaluated in usability studies of audiovisual maps (Chapter 6 and Appendix B) and the application called *Canada-USA Commodity Trade 1976–2000*, from which this example was taken, will be further discussed in Chapter 8 and Appendix H. As will be discussed more fully in Chapter 6, these dynamic audiovisual legends seem to be useful but at least some users have required guidance when first working with them.

(R5: ✓) Listing a set of parameters for a specified set of areas or features is an explicit use of a coordinated, generally tabular, display of additional data related to mapped features. As such, the use of coordinated auditory representations can be thought of as analogous to this in much the same way as discussed above for adding and removing layers (R3). The idea of listing parameters for a subset of features, however, emphasizes a more automated approach to the subsequent examination than the interactive re-expression approach suggested above. This suggests instead that the specified set of parameters be played according to a script for all of the specified features, with each feature being represented one at a time by the sonic representation of its parameters and all features presented in some understandable order, such as: alphabetic by feature identifier, temporal order (if appropriate), or ordered according to an index derived by sorting the values for one of the represented parameters. Visual feedback indicating the feature(s) currently being represented would be essential to help a user understand the current context for the auditory representation being played and requires synchronization of the visual map and linked auditory elements. The set of features to be included in this sequential ordering could be, for example, explicitly selected by the user (possibly, but not necessarily, through the use of cartographic zooming) or could be comprised of all mapped features. The application *Canada-USA Commodity Trade 1976–2000* uses this approach when sequencing a set of map views selected according to user specified criteria, synchronizing visual map updates and linked auditory representations of related variables.

(R6: ✗) Although listing all or additional parameters for *one* specified area or feature is not that different from listing the parameters for all features, in an auditory representation it is debatable how widely useful the presentation of one set of observations would be (except if read out using synthesized voice or narration). For many people (those without the ability to detect and recognize absolute pitch, for example), the auditory encoding of data yields a representation that is most useful for discerning differences or patterns rather than for ascertaining absolute values (Kramer 1994b). Perhaps with sufficient training or a certain amount of innate listening ability, a person could discern information by listening to just the parameters associated with one feature but it is more likely to be useful as part of a sequence presenting the parameters for multiple features.

(R7: ✓) Comparing observed values to statistically significant values is an additional application of using sound to detect differences within a data set. This comparison, for example, could be conducted over time by playing the representation for an observation preceded or followed by the representation of a data set mean. But this comparison could also be performed, at least for one or two variables, by simultaneously playing the statistical value as a separate sound encoding along with the presentation of a specific observation. This technique could also be used in conjunction with a presentation of sequenced observations; the statistical mean, for example, providing a steady accompaniment to each observation throughout the sequence.

(R8: ✓) As will be discussed in more detail with respect to specific audiovisual map examples in Chapters 7 and 8, there are a variety of design options for the use of abstract sound as part of audiovisual cartography. Some representations will be more effective than others for presenting certain types of relationships but this may also be influenced by individual listening preferences and experience. Thus it would be ideal for users and designers of an audiovisual map presentation to be able to experiment with alternative sound designs (detailing the content of the sound as well as the manner in which user interactions modify those sounds). This is the auditory equivalent of modifying the graphic symbolization of a map. Considerations concerning the distribution of the underlying data and the implications of that distribution for the choice of a classification method potentially apply equally to auditory and graphic data representations.

The design and implementation of toolkits for the creation and configuration of sound designs appropriate to widespread audiovisual cartography remains a significant future effort. Computer audio subsystems vary widely and the range of loudspeakers and headphones makes consistent audio reproduction and representation across this diversity very difficult.

Canada's Trade with World Regions: 1976–2000 (Chapter 7 and Appendix B) includes user interface controls that allow a user to select from among a predefined set of sound designs for the representation of two linked variables.

(R9: ≈) Emphasizing particular data through blinking does not have an obvious auditory analogy when considered with respect to the two dimensional role that Monmonier (1992) proposed for visual blinking, although spatialized audio (even just using left-right stereo panning) could potentially be used, in coordination with vision, to draw a user's attention to an area of a map or to particular features on a map. In addition, based on the metaphor that sounds are caused by events, temporal placement of audio alerts could be used to draw a user's attention to a particular data observation in conjunction with the sequenced playback of a number of observations (see R5 above; Hermann et al. 2003).

5.2.2.2 Interactions with Data

(D1: ✓) Knowledge Discovery from Databases (KDD) aims to develop an understanding of information stored and encoded in large repositories of data and to find emergent patterns in that data for purposes of hypotheses generation and anomaly detection. Miller (2007) divides the process into the following major steps: understanding the application domain; data processing, generalization, and reduction; data mining; and interpretation and knowledge construction. Because of the focus on pattern detection and hypotheses generation prior to steps that may validate patterns and hypotheses, KDD is iterative, interactive, and requires the use of tools that assist in analyzing intermediate results and formulating successive operations, still guided by "a high-level, most likely human, intelligence" (Miller and Han 2009, 9). There remains a tendency in the literature to associate the value of human intelligence with the human visual system, as demonstrated by Miller and Han continuing to argue that "[t]he human visual system is extremely effective at recognizing patterns, trends, and anomalies", capabilities that "can be exploited in many stages of the KDD process, including OLAP [online analytical processing], query formation, technique selection, and interpretation of results" (ibid.). Positing that a visual bias concerning characterizations of human intelligence has been extended to biases in the development of tools supporting that intelligence, it is possible to argue that current tools and methods favour visual processing but don't necessarily preclude a research agenda that enhances auditory processing and methods as complementary to current visual ones. Indeed, despite the above characterization of the stages of KDD, querying data, interpreting results, and analysis could all equally be argued to be abstract, logical, and/or mathematical processes. At least some of the results from these processes, as has been argued by human-computer interaction researchers, may be amenable to auditory representation and manipulation as well as visual (Flowers 2005). As is argued throughout this thesis, the use of sound as part of processes such as this has not yet been sufficiently researched.

(D2: ✗) As with some of the steps in the KDD processes described above, logically combining existing spatial layers or the attributes of an existing layer or generating new attributes or layers from existing data are common and necessary steps in preparing and processing data for geographical analysis. These operations tend to be logical, mathematical, or more generally rule based operations that, as argued above, are not intrinsically visual. As discussed above for the results of intermediate and final processing steps in the KDD process, the use of audiovisual cartographic techniques, if sufficiently flexible and easy to define, would be appropriate as part of the process (but that is covered in D1).

(D3: ✗) Inspection methods such as cartographic cross-classification or brushing applied to geographical data (Monmonier 1989), relying on interactions with multiple abstract two-dimensional representations of data observations, and logical relations between representations, have no obvious auditory analogies. Research has suggested that auditory displays can be effective in conveying information concerning correlation of a single pair of variables in the form of a scatterplot (Flowers et al. 1997, 2005). Thus possibilities exist for integrating linked auditory representation as part of a strategy to analyze the brushed data but procedures for applying the 'brush' are currently understood as predominantly visual.

5.2.2.3 Contextualizing Interactions

(C1: ‡) Linking sound representation to spatial data is here proposed as one of the ways in which sound can be productively used in audiovisual cartography. Coordinated design of visual and acoustic elements, fulfilling the functions assigned to dynamic audiovisual legends in this thesis at least during a period in which a user is learning to work with map representation, is required to ensure this is a useful and usable addition to the techniques of interactive cartography.

(C2: ✓) The technique of displaying multiple, coordinated, thematic maps is similar to the simultaneous representation of multiple variables using linked audiovisual representation, as outlined for (R3) above, except for the consideration of how the currently sonified spatial extent relates to the visually mapped extent. As discussed above for (R1), the sonified extent could equal, exceed, or be specified within the visually mapped space but the means of selecting this extent and of signifying to the user what extent has been selected must be clear. The ability of users to relate the sounds and the visual elements to which those sounds apply is obviously important and any inability to perform this feat would undermine the approach. But this ability on the part of interactive map users has been demonstrated by other research (Fisher 1994; Krygier 1994; see also Chapter 6).

(C3: ✗) The technique from animated cartography of displaying a sequence of images showing changing cartographic representations, and thereby offering different perspectives on mapped patterns, may not have a useful auditory analogy. Although it would be possible to present multiple auditory representations of the same data, the greatest potential of the types of non-speech sonification being discussed here are in the abilities of listeners to discern pattern, continuity, change, and rhythm within a representation over time (Kramer 1994b). Changing the representation as well then creates multiple possible reasons for any acoustic changes detected by listeners and attendant uncertainty by them concerning the intended meaning.

(C4: ✓) Dynamic comparison of a specified subset or all observations in a defined order seems like a very promising adaptation of graphic scripts (Monmonier 1992) and the auditory equivalent of visual re-expression (DiBiase et al. 1992) based on the metaphor that auditory change indicates phenomenal change (again, taking the potential for differences in the represented visual and auditory extents into account). There are two approaches to this that would appear very useful and would be relatively easy to implement using straightforward algorithmic scripting: a) temporally ordered sweeps of timestamped observations of a set of attributes for a single feature or region; and b) re-expressions of a set of variable observations, each 'frame' of the re-expressed sequence sonifying the data for one of a set of mapped features or regions, based on sorting a selected attribute by value.[8]

[8]The ordering of observations may be re-expressed according to almost any criteria (with respect to visual animation, see DiBiase et al. 1992). But note the concerns of Tversky et al. (2002, 258) who argued that "there should be a natural correspondence between change over time, the core of animation, and the essential conceptual information to be conveyed [...] be [it] change over time, [...] temporal sequence, or causal flow." Careful design and user evaluations of both visual and

Alternatively, it may be useful to allow a user to define an arbitrary ordering based on current and past examination of a data set, allowing them more flexibility to pick and compare observations, if desired. For example, selecting points at intervals along a marathon route and sounding out a representation of the associated elevation values could be a quick means of evaluating the appropriateness of the terrain for the race. Similarly, selecting a pair of observations and specifying that the auditory sequence should be played in a loop would create the auditory equivalent of flickering (Monmonier 1992): sonic quasi-movement now indicating change. Monmonier also suggested alternative pacing of visual animations which would be applicable to the use of auditory representations. Ordered observations could be sequenced with uniform pacing for each observation or, to highlight clustering within the observations, pacing for each according to a linear sweep of the observed value range. The use of re-expressed and temporal audiovisual sequencing in *Canada-USA Commodity Trade 1976–2000* and some implications of the different possibilities for sound design afforded by configured sequencing as opposed to interactive selection are discussed further in Chapter 8.

(C5: ≈) The dynamic use of successive build-up, although not serving the same purpose in auditory representation, could be quite useful as a means to allow a user to examine certain variables or subsets of the auditory encoding in isolation as a method for learning the auditory encodings or to focus on that subset during analysis. A dynamic audiovisual legend, as discussed above, provides one means by which a user can listen separately to each auditory representation during training. The addition of per-variable gain or mute controls to the user interface of an audiovisual map application, as will be discussed in Section 8.1 with respect to

auditory re-expression are required to ensure that users are able to learn and understand the basis of comparison and ordering.

Canada-USA Commodity Trade 1976–2000, allows a user to focus on a subset of the defined auditory representations as desired.

(C6: ✓) Spatialized sound could play a role in conjunction with three-dimensional navigation or fly-throughs and has been studied as a component, along with simple two-dimensional map-like graphics, of educational virtual reality systems (Sánchez et al. 2002). Adopting a first-person viewpoint, such an approach is very similar to navigation within three-dimensional games. Lack of standardized media formats for spatialized audio on the World Wide Web currently limit the use of this approach for browser-based applications.[9] Without using spatialized audio, Coburn and Smith (2005) generated synthesized music from satellite imagery, deriving the data driving the acoustic synthesis, and thus the spatial focus of the sonification over the duration of the performance from predefined (non-interactive) transects across the terrain. They argued that different terrains (e.g., urban, residential, parkland) generated distinct rhythms in the sonifications. Shilling and Shinn-Cunningham (2002) argued that, in virtual environments and simulations, sound has been found to increase users' engagement with subject matter and to increase their emotional response to modelled situations.

5.2.2.4 Interactions with Temporal Dimension

(T1: ✓) Similar to the discussion above of animated re-expression with audiovisual representation (C4), the use of dynamic comparison using temporal order is a very promising use of coordinated audiovisual representation and has been used as part of this research in *Canada-USA Commodity Trade 1976–2000* as discussed in

[9]Software supporting acoustic spatialization exists and could be used as part of games or web applications but would currently need to be installed as a separate component on each system accessing the map, limiting is usefulness in the broad dissemination of maps. The lack of standard media formats for spatialized sound data would further complicate the preparation of audiovisual map data and representations.

Appendix H. Testing concerning pacing and the use of visual feedback to ensure a user understands the context of the representation are still required.

(T2: ✓) The manipulation of dynamic variables such as duration and period for the auditory representation of sequenced observations could be used, as in visual display (DiBiase et al. 1992), to stress, for example, periods of high activity. Shilling and Shinn-Cunningham (2002) argued that human hearing can discern higher-resolution temporal patterns (rhythms) in sound, and this capability can be used to complement the perceptions of our visual system. Shams et al. (2004) argued that sound has been shown to be capable of altering, for example, the temporal ordering of visually perceived stimuli, the number of perceived visual events, and the perceived location of visual events. Therefore, there may be opportunities to use auditory representation on a different time scale from that used for visual map sequencing, perhaps running finer grained sequences of sonification within coarser scale visual animations, but task appropriateness would need to be considered and such a design would need careful testing.

(T3: ✓) Interactive controls for audiovisual representations, providing configuration and activation controls to adjust overall pacing, duration, visual and temporal extent, certainly have a role to play in the user interface. In conjunction with the possibility, as discussed above for (T2), that there may be opportunities to consider different time-scales for the visual and acoustic representation, there may be additional possibilities to separately stretch or compress auditory playback rates or the configuration of the range of acoustically-represented observations within each visual step. Careful design and testing would be required.

—

Although this section has focused on abstract applications of non-speech sound, the analysis should be understood within the framework concerning the creation of meaning through the use of sound as part of audiovisual cartography discussed in Section 3.1 and depicted in Figure 3.1. Importantly, although this discussion has focused predominantly on sound as a medium conveying information in conjunction with a dynamic visual map, such uses do not preclude possibilities for shaping meaning within the map's visual signification and do not preclude other possible motivations for the use of sound including the establishment of mood, encouraging user engagement, and thematically linking a map to other materials through shared sound design, depending on application objectives and audience. The use of linked audio also does not preclude other potential uses of auditory media including, but not limited to, narrations providing background information and application usage overviews, or contextual audio referencing themes, peoples, or places highlighted in related materials.

Any of these elements chosen for use in an audiovisual map application need to be considered in relation to each other, the visual design, the behaviour of the application, and with respect to how each is presented within the application and controlled or varied by user actions. User interpretation should be expected to emerge from their encounter with the entire matrix of materials, and from training or guidance provided concerning overall design intentions. The state of knowledge concerning audiovisual cartography is limited and many questions remain for future research.

Part III

Designing and Evaluating Audiovisual Maps

Earlier parts of this thesis discussed the contemporary research context for audiovisual maps to be distributed and used over digital networks, and motivations and methods for the use of sound in a range of disciplines that could inform audiovisual cartography. Based on those analyses, this part of the thesis discusses user responses to a set of audiovisual map examples, technological solutions enabling the use and distribution of audiovisual maps over digital networks, and highlights some design lessons learned through the process.

User evaluations were conducted using map examples developed using an earlier version of the cartographic sound subsystem to be presented in this thesis. Lessons learned, through those evaluations and the creation of the examples used, contributed to the redesign of that system. Therefore, this part begins with a discussion of the user evaluations in Chapter 6. Chapter 7 describes the sound subsystem design. Chapter 8 provides an overview of *Canada-USA Commodity Trade 1976–2000*, an example developed along with the audio software rewrite, and discusses some audiovisual design issues considered during the development of this example. The final chapter summarizes the results of this research and discusses possible future research.

6

User Evaluations of Audiovisual Maps

This chapter presents an overview of user-centred design (UCD), and usability tests within that process, before discussing the methodology used and an analysis of the data generated in usability tests conducted with university students asked to work with and comment on four audiovisual map examples. The chapter concludes with a summary of the test results, a discussion of the testing process, and suggestions for additional user evaluations that could highlight types of issues other than those that are the focus of the testing reported here.

The approach to evaluating the selected audiovisual maps was based on usability testing methods developed as part of UCD methodologies that have been applied to design problems for a broad variety of systems: information systems, machines, appliances, and buildings (Baecker 1995; Ylirisku and Buur 2007). Nielsen (1993, 24–25) argued that a system is *useful* if it provides *utility* in accomplishing some purpose or set of tasks and if it is *usable*, meaning that users can reasonably easily take advantage of the system's utility. Usability testing is then an exercise in which someone thought to be representative of potential users of a system is asked to perform certain tasks using the system under test, with data generated through the testing process that can later be analyzed to understand users' reactions to working with the system, and to identify and understand when and why users had difficulties.

User-centred design approaches for evaluating computer system interfaces have been adopted by designers of interactive maps (Kramers 2009; Schobesberger 2009; van Elzakker and Wealands 2007). Two projects on which I worked during early phases of this thesis research as a member of the Carleton University Geomatics and Cartographic Research Centre (GCRC), in coordination in particular with human-computer interaction (HCI) researchers from the Carleton University Human Oriented Technology Lab (HOTLab), were structured around processes that incorporated incremental design and usability evaluation, applied iteratively, to create cybercartographic atlases. The first being the *Cybercartographic Atlas of Antarctica* (Pulsifer et al. 2007, 2005), from which two of the evaluated audiovisual map examples discussed in this chapter were derived. The second being the *Cybercartographic Atlas of Canada's Trade with the World* (Eddy and Taylor 2005), which provided data I used to develop one map example included in the evaluations.

Although the user evaluations discussed in this chapter were conducted separately from earlier evaluations done for these two GCRC atlas projects, those earlier UCD processes provided contextual information that was carried over into the audiovisual map examples described here and the design of the evaluation processes. Therefore, before detailing the results of user evaluations of audiovisual map examples, Section 6.1 discusses UCD and sampling issues for usability testing as part of UCD. Section 6.2 briefly outlines some of the decisions made through the UCD processes for those atlases that were incorporated into the design of the map examples included in these evaluations; and discusses the methods used for the user evaluations, including how the evaluation sessions were conducted, the data generated during each session, and how the data from these sessions have been analyzed. Section 6.3 discusses what the data indicate about each of these audiovisual map examples. For reference, Appendix B describes the map examples used during the evaluation sessions and video captures included on the CD-ROM accompanying this thesis show these map examples being used (see Appendix G).

6.1 User-Centred Design: Background

UCD processes emphasize the identification of user groups that would have an interest in using the outcome of a design process,[1] the analysis of requirements to be met to ensure that the resulting outcomes are usable by members of the identified groups and in the context in which they would be expected to use them, and a set of tasks these users must perform or scenarios in which the outcome of the design process would be used. Results from the user focused analysis become central pieces of the design specification, subsequently refined through iterative development processes that guide design decisions and from which testing scenarios can be derived (Baecker 1995; Schobesberger 2009; van Elzakker and Wealands 2007; Ylirisku and Buur 2007).

The objective of UCD and of user evaluations conducted as part of the process is understanding the characteristics of the system being designed, how users want to work with it, what hinders their efforts, and what additions or changes to the system would improve their experience in working with it. Certain proposals for UCD methodologies or applications of the methodology, such as that by van Elzakker and Wealands (2007, 491), include strong statements concerning the importance of *knowing your users*. Despite such statements, the objective of understanding system users in UCD is, somewhat more narrowly, focused on learn-

[1] The outcome of a design process could be a product, service, system or, more broadly, some *change* (Ylirisku and Buur 2007, 8).

ing why, how, where, and when they carry out a certain task or use a particular system and to what end. This then informs system design decisions; the selection of participants for testing based on criteria including expected skills, needs, and characteristics (e.g., age, gender, socioeconomic factors); and the definition of tasks to be included in testing. The rationale for user evaluations as part of UCD is oriented toward pragmatic and economic considerations that value fast cycles of project-specific testing, with findings from those tests used as quickly as possible to update project designs and to continue the iterative process. The desire to conduct user testing within an overall iterative process often requires that you test what the system can currently do, or what can be adequately emulated, *with evaluation participants you are able to recruit* who are reasonably similar to the target audience identified during requirements analyses. Thus, while imperfect, the alternative may be an inability to conduct any user testing.

Proponents of usability testing methods, as part of iterative UCD processes, recommend participant sample sizes that seem remarkably low to many researchers with experience of surveys, interviews, or observation of human subjects in other disciplines. Nielsen and Landauer (1993) argued that if multiple representative users repeat usability tests independently, then the probability of discovering any specific problem during those tests is independent of whether or not that problem was found in previous tests. Although the actual problem discovery rate may vary significantly between testing scenarios (e.g., tests of different systems or different scenarios with the same system), the new problem arrival rate can be monitored during testing and factored into a cost-benefit analysis to determine when to stop testing. Assuming that there are a fixed number of problems to be found,[2] Nielsen

[2]cf. Macefield (2009) who argued that problem identification and classification is far more difficult and subjective than such a formulation would indicate. In addition, there is no way to know whether or not all problems, however defined, have been found.

and Landauer argued, continued testing will suffer diminishing returns as measured by the number of new problems detected in successive test sessions. Virzi (1992) argued that conducting test sessions with as few as four or five test participants was sufficient to find 80% of all usability problems.[3] Macefield (2009) argued that sample size decisions in usability studies depend on acceptable costs, the type of study (distinguishing, in particular, between problem discovery and comparative studies), and maturity of the system being tested. He argued that 3–20 participants is typically sufficient and that for problem discovery, especially with novel interface designs, 5–10 is often enough.

Lindgaard and Chattratichart (2007), by contrast, examined the results of independent usability tests conducted with the same system and argued that there was no correlation between numbers of test users and either the number of new problems or total number of problems found. Instead, Lindgaard and Chattratichart argued there was a correlation between the number of unique test tasks performed by users and the numbers of problems found indicating that system function coverage may be a more important criterion for effective problem discovery. The representativeness of participants, compared to the population of actual or potential users of a system, can impact the effectiveness of usability tests and, as argued by Lindgaard and Chattratichart (2007, 1423), poor recruitment practices (e.g., too many similar participants, allowing expert users to participate) can reduce the effect of good task coverage.

Whereas arguments concerning the number of participants rely on the probability of users finding specific problems within a system, the methods of data generation used during user evaluations can provide materials suitable for both

[3] Lewis (1994) argued that examination of usability results from an independent study corroborated the conclusions of Virzi.

qualitative and quantitative analysis of the performance and behaviour of the system under study and of the responses of participants to that system. The logic of sampling for quantitative and qualitative research methods ideally can be different with the objective of quantitative sampling being representativeness and the objective of qualitative sampling often being contextual theory development or purposeful selection to answer certain questions with sampling, again ideally, continuing until no new concepts emerge from additional cases (Sandelowski 1995; Corbin and Strauss 2008, 143–157; Mason 2002, 120–144). Within the context of UCD processes, user evaluations may be guided by expectations of the relationship between the number of participants and the probability of finding new problems but, in practice, test practitioners stop testing when ongoing test sessions cease uncovering new problems, even though there is no way to be sure other problems would not be found if sessions continued (Faulkner 2003).

6.2 Usability Tests for Audiovisual Map Examples

6.2.1 Context for Evaluations

The original user needs analyses conducted for the *Cybercartographic Atlas of Antarctica* (CAA) and the *Cybercartographic Atlas of Canada's Trade with the World* (CACTW) identified high school students and subject matter experts (e.g., scientists working on Antarctic research, in the case of the CAA) as distinct user groups and recognized that these groups would have distinct requirements (Pulsifer et al. 2005; Rasouli et al. 2004). Initial user requirements analyses for these projects included stakeholder meetings to develop long-term visions for meeting the requirements of all identified groups but project scoping eventually identified user needs of the student groups as a minimum objective for the projects and project activity was focused more on those requirements. User requirements analysis then included interviews with educators and reviews of Ontario high school curricula to understand if the subject matter for each atlas would potentially relate to educational materials available to students and possible student project topics (Baulch et al. 2005; Eddy and Taylor 2005; Pulsifer et al. 2005).

Iterative design, prototyping of user interfaces for the atlases, and usability testing were conducted based on then most-current results of requirements analysis and earlier testing, although university students were often brought in as test participants in place of high school students because they were easier to recruit.[4] In addition, definition of specific task-oriented requirements derived from the examinations of curricula has been difficult because the potential breadth of subject matter that could be incorporated into these atlases and potential inter-relations among those materials are innumerable but individuals or groups tasked with any specific school project would probably only rely on either of the proposed atlases for a part of the information needs of their particular focus (Eddy and Taylor 2005).

Within the cybercartographic atlas projects, user evaluations were conducted to assess information organization alternatives, visual layout design alternatives, and user interaction models. No user evaluation tests examined the use of sound as part of the atlases because they were all completed prior to the development of sound design and reproduction capabilities within the software used to create these atlases.

[4]Ethics approval for user testing involving high school students is, in theory, not markedly more difficult than testing involving university students. Obtaining permission from Ontario school boards to go into schools to conduct tests or to contact schools to recruit participants has, in practice, become so difficult that Carleton University's ethics review board now discourages student-researchers from trying. In addition, the Carleton University Department of Psychology, members of which participated in conducting the evaluation session described here, has established processes that significantly ease the task of recruiting undergraduate participants.

6.2.2 Usability Testing Overview

Once the functional capability to incorporate sound as part of map designs had been developed in the atlas software and a number of map examples had been developed using sound, arrangements were made to conduct usability tests in conjunction with researchers from the Carleton University HOTLab.[5] During October and November 2007, students enrolled in Psychology 1001 and 1002 courses at Carleton University were recruited to evaluate four audiovisual map examples in exchange for credit toward their course grade. Students in these courses were thought to represent an appropriate evaluation population for the planned test design. They are generally computer literate and familiar with web browsing but would be unlikely to have prior knowledge of audiovisual maps.

The four audiovisual map examples, described in detail in Appendix B, use a variety of types of sounds (speech, music, sound effects), were designed to use the sounds for different purposes, initiate sound playback in response to user actions over the map (or parts of the map) and sometimes in response to user actions related to other interface components (audiovisual dynamic legends, text hyperlinks), and provide differing levels of explanation within the interface concerning the use of sound and the purpose for the use of sound. The use of these four maps

---

[5] This project was funded in part under a Social Sciences and Humanities Research Council (SSHRC) Image, Text, Sound, and Technology (ITST) Research Grant. Dr. D. R. Fraser Taylor was the Principal Investigator, Dr. Sébastien Caquard collaborated on the research. Dr. Avi Parush sponsored the ethics application to the Carleton University Department of Psychology Ethics Review Board, and supervised (then) graduate students Shelly Roberts and Aren Hunter, who jointly conducted the evaluation sessions. I spearheaded the writing of the grant proposal, created two of the map examples used during the evaluations, contributed to the map examples taken from the *Cybercartographic Atlas of Antarctica* (Caquard et al. 2008), and wrote the ethics application package and evaluation scripts in conjunction with the HOTLab researchers (Roberts, Hunter, and Parush). Although I attended one of the evaluation sessions as an observer, I mainly stayed away to avoid my personal investment in the development of the map examples and technology from potentially influencing the tests. Roberts et al. (2008) reported the results of testing along with recommendations for possible modifications to each of the tested map examples.

was intended to provide test material that would generate data concerning specific usability problems with each audiovisual map application, and user responses to different uses of sound as part of these applications.

Participants were recruited and each worked with three of the audiovisual map examples.[6] After completing sessions with eight participants, resulting in a total of 24 map evaluation sessions and each map being evaluated by six participants, the HOTLab researchers thought continued sessions were unlikely to produce new feedback concerning the examples and we stopped conducting sessions. Table 6.1 provides an overview of the audiovisual map examples tested during the user evaluation sessions showing, for each application and user, the numbers of user evaluation sessions conducted, the order in which each user was presented with each of the audiovisual map examples, and a reference to the section of Appendix B containing an overview of the application and discussion of how that application was presented and used during the sessions. The ordering information provided in Table 6.1 details both the order in which each participant tested the three maps shown to them and the design done to ensure that session participants would not all see the same subset of map examples nor the same ordering of examples.

The maps that would be evaluated by each participant were predetermined according to the order in which the evaluation sessions were conducted. No attempt was made to balance the numbers of participants that worked with each map according to any particular shared characteristic such as gender.[7]

[6]One pilot was conducted with a volunteer graduate student evaluating all four of the map examples, causing us to refine the test scripts and to decide that we would ask each participant to work with only three of the audiovisual map examples to ensure that the sessions for one participant could be completed within one hour.

[7] As it happened, an equal number of males and females worked with two applications, four males and two females worked with one example, and two males and four females worked with the remaining map.

Table 6.1: User evaluations: Map examples testing summary.

Audiovisual Map Example	Map Test Order by User†								Test Totals
	1(F)	2(F)	3(F)	4(M)	5(M)	6(M)	7(M)	8(F)	
Ottawa Area Federal Election Sound Map: June 28, 2004 (OAFESM: §B.1)	3		2	1	3	2		1	6
Canada's Trade with World Regions: 1976–2000 (CTWR: §B.2)	2	1	3			1	3	2	6
Antarctic Territorial Claims (ATC: §B.3.1)		2	1	3	2	3	1		6
Antarctic Explorations (AE: §B.3.2)	1	3		2	1		2	3	6
Total									**24**

†Column header is participant number (the order in which they participated in the evaluations) with the participant's gender in parentheses. Each column shows the order of map application evaluation within a single user's session. Each row across shows at which positions within the group of three evaluations performed by each user the map application was tested.

6.2.3 Data Generation

Considering that these were to be the first evaluations of users' responses to audiovisual map examples, by me or within the project, methods that generate data appropriate to qualitative analyses of users' evaluations, understandings, and responses to the examples seemed most appropriate. This did not necessarily preclude the collection of quantitative data or quantitative analyses of data generated during the sessions but rather reflects the judgement that it would be premature to focus only on metrics concerning user performance prior to understanding whether or not users understand these audiovisual maps and how to work with them. Suchan and Brewer (2000) argued that increased use of geographic visualization interfaces, encouraging private map creation and use as a means to examine unknowns, have caused an increase in the use of qualitative methods by cartographic researchers along with quantitative methods. In particular they noted the ability of researchers using qualitative methods to "generate and then modify initial research conceptions; it helps see what events lead to what outcomes or consequences; and it leads to new integrations" (ibid., 147). Olson (2009, 1) argued that "[m]ultiple methods (quantitative and qualitative) are necessary for understanding human interaction with maps" (also Schobesberger 2009). van Elzakker and Wealands (2007) argued that "[i]n many cases we do not yet know enough about how maps or Multimedia Cartography products work for knowledge construction or geospatial problem-solving and, therefore, it is difficult to formulate hypotheses for testing by means of quantitative techniques."

Very little empirical data exists concerning the abilities of users to understand and work with auditory representation incorporated into maps. Therefore, these tests were designed to generate data that would allow an analysis of users' abilities to work with the maps during a first encounter and to enable a close inspection of causal links for user behaviour in response to the maps. To that end, the main focus of the evaluation sessions was the generation of qualitative data and the highest priority during planning for the sessions was the desire to be able to conduct qualitative analysis of the data once generated.

The four examples evaluated were: *Ottawa Area Federal Election Sound Map: June 28, 2004* (OAFESM), *Canada's Trade with World Regions: 1976–2000* (CTWR), *Antarctic Territorial Claims* (ATC), and *Antarctic Explorations* (AE). Each has a different purpose and sound is used differently across the set of examples (Appendix B).

This range of purposes, designs, and explanation concerning the use of sound was thought to provide a good basis for an examination, in addition to questions concerning the usability of each application, of at least the following set of questions:

1. Are users able to work with the audiovisual maps?
2. Do users understand the audiovisual maps, including the relations between

sound and visual interface elements?

3. Are users' understandings of the maps modified by the use of sound?
4. How do users react to the audiovisual maps?

The above questions broadly attempt to examine the usability and usefulness of each of the presented audiovisual map examples. Question 1 addresses the usability of the audiovisual map applications, which includes many issues concerning the design such as adequacy of explanatory materials describing the intended purpose and use of the application, or appropriateness of the user interface which could in turn be based partly on familiarity of users with similar user interfaces. Because of the paucity of sound use in maps and the limited available information concerning users' responses to audiovisual maps, usability during initial map encounters may be an important factor in users' future willingness to engage with such artefacts.

Questions 2 and 3 broadly examine the usefulness of the audiovisual map examples by assessing whether or not users understand the information presented, including sound as one of the signifying modalities. An additional goal of the evaluations, because of the particular focus of this research, was to determine whether or not there are indications that the use of sound as part of the application influenced users' understandings of the information presented.

Question 4, in part, attempts to generate data concerning user acceptance of each audiovisual map example: do they like the application and does it seem to be something they are comfortable using? This question is related to usability and usefulness because research has indicated that a user's mood may have an effect on both qualitative and quantitative measurements of task performance, and a person's affective responses may both be influenced by and may influence estimations of satisfaction and acceptance.[8] But this question also potentially examines emotional responses by a user to the subject matter presented in a map example. In addition, this question could generate data concerning whether or not a user recognizes emotional characteristics of the subject matter in an audiovisual map example or any role that sound may play in communicating indications of emotion. The relations between emotion and sound in other media[9] indicate that this may require consideration.

Studying usability of these examples during repeat sessions would significantly alter the circumstances for basic usability questions such as whether or not a user found the sounds in the first place or understood the basics of the audiovisual design.[10] Consistent with the objective to examine initial encounter usability, map examples were presented to students with relatively little initial explanation of

[8]Erevelles (1998, 199) defined *affect* as a "valenced feeling state" and *mood* as a special case of affect that "is relatively low in intensity, and is usually unassociated with a stimulus object." Hirt et al. (1996) argued that mood was a factor in both task interest and in task performance but also that the causal relationships between mood, interest, and performance are complicated. Because their test was designed to induce mood (happy, sad, or neutral) prior to task performance measurement, it is not clear that their results directly apply to the short duration evaluation tasks used in the evaluation sessions discussed here. In marketing research, customer satisfaction (or dissatisfaction) responses are theorized as having both cognitive and affective components (Erevelles 1998; Oliver 1993). Brave and Nass (2003) argued that emotion should receive greater consideration in HCI research because emotions have potential effects on a user's attention, memory, performance, and assessment. Agarwal and Meyer (2009) argued that emotion has been largely overlooked within HCI research, despite a strong relationship between emotion and overall user experience, possibly due to complexities in assessing emotions (cf. Norman 2004; Picard 1997, 1998).

[9]Sections 4.1 and 4.2 review potential effects of sound on emotion in cinema and games, respectively.

[10]Repeat sessions with each user would provide opportunities for other types of questions to be considered. Nevertheless, because part of the goal of these evaluations was understanding the usability of these audiovisual maps, it seemed reasonable to assume that problems could be found that would require attention before effective longer term use studies could begin. Attempting to conduct repeated use testing is probably best done using a map example that is already known to be relatively usable and that provides enough richness in subject matter to allow repeat sessions to be conducted using new tasks supported by the application. Because this was the first set of evaluations for all of these audiovisual map examples, the level of usability for each was *not* known and, at least for some of the prototypes, it was not clear that they provided a rich enough set of tasks to warrant repeated sessions. Research including repeated test sessions with individual users would be useful and remains a future possibility.

how to work with the map or of the purpose of the designs (see evaluation session scripts in Appendix D).

As is common in user evaluations that are part of UCD processes, the evaluation sessions were conducted by HCI researchers rather than by the people who had designed the map applications and I, with the exception of acting as an observer during one session, was not present.[11] For the evaluation sessions, scripts and questions were designed to *directly* generate data about the usability and usefulness of the examples for each user. For each audiovisual map example, the interviewer would open it in a web browsing application window, provide the participant with a general scenario, and ask them to familiarize themselves with the example. If a student appeared able to understand the example, then they would be asked to answer specific questions using the map. If they had difficulty working with an example, their difficulties were to be noted for later analysis, they were to be given assistance to ensure that they understood how to work with the example and, when comfortable, they would move on to the directed questions. Finally, each participant was asked to complete a satisfaction questionnaire concerning the example.[12]

Each participant was asked to *think aloud* as they worked with each map example to generate what Lewis and Rieman (1994, 5.5) referred to as *process data*, "observations of what the test users are doing and thinking as they work with the tasks." Periodically the interviewer would ask the participant to clarify what they were thinking about or looking for as they worked with an example, especially if they seemed unsure how to proceed at a particular time during the evaluation.[13]

[11]See footnote 5 on p.145.

[12]Evaluation scripts and satisfaction questionnaires are presented in Appendix D.

[13]This usability testing method is also known as *concurrent verbalization* (Ericsson and Simon 1993). Ericsson and Simon developed the method with the intention of having participants in evaluations provide a running verbalization of mental *actions* (e.g., 'reading', 'searching for') and

The user evaluation sessions include some direct questioning by the interviewer concerning whether or not a user liked or disliked a map example or the audio used in it, sometimes cued by the interviewer's interpretation of a user's reactions during a session but also explicitly as part of the satisfaction questionnaire administered after a participant worked with each example. Beyond questions concerning a user's level of comfort with the application as a whole or the use of sound in it, other data related to the emotional responses of users or of their interpretation of the map examples as containing emotional content emerged from the less structured discussions between the participant and interviewer or from comments made by the participant as part of their concurrent verbalizations.

During the evaluation sessions the following data were collected for each map example presented to each participant: *observation notes*, an *audiovisual screen capture*, and a *completed satisfaction questionnaire*. The interview team, the interviewer and an observer, recorded observations concerning each participant's experience with each map example, noting: whether or not a user had difficulty using a map example; whether or not assistance was given to help a participant with an example; relevant comments made by the participant while working with an example; answers to clarifying questions asked during the evaluation; and the participant's answers to the scripted questions. During each evaluation session, audiovisual screen capture and audio recording software was used to create a video recording of the computer screen showing the map example, the user's cursor location on the screen, and the application's screen updates as the user worked with it. In addition, the audio track of the video recording captured the computer system's audio, including the sound track of the map example, and any conversations that

---

the *objects* on which those actions are performed (e.g., 'text', 'map', 'narration'). Lewis and Rieman (1994) also proposed the use of the *thinking aloud method* as part of a distinction between *bottom-line data*, summary data that explains what happened, and *process data*.

occurred during the session (discussion between interviewer and participant and the participant's concurrent verbalizations). Finally, each participant completed a questionnaire concerning the usability and usefulness of the map example, asking what they liked least and most about each application, and asking for suggestions they felt would improve the application.

6.2.4 Data Analysis

The analysis presented in Section 6.3 is based primarily on a cross-sectional analysis (Mason 2002, 150–165) of transcriptions of the audiovisual computer screen recordings, relying on participant verbalizations and the discussions between participant and interviewer concerning their interactions with each application more than their retrospective opinions concerning the application as captured in the satisfaction questionnaires. Although I consulted the questionnaires during analysis, these were not my main focus. The transcriptions and video recordings provided more information concerning users' immediate reactions to application behaviours because of the ability to concurrently see and hear what the application is doing, hear what the participant and interviewer are saying, and to see where the participant has placed the computer cursor on the screen. Tracking the location of the computer cursor on the screen can often explain not only why the application is behaving as it is but, when a participant is using an interactive map application, it can also often show approximately where in the application that participant's attention is currently directed (e.g., participants often talk about what they are doing and concurrently select blocks of text or place the cursor on the object they are talking about). This can be very useful in trying to decide to what a participant is responding when they make a particular comment, although this is also a time consuming approach to analysis.

Of the four questions outlined in Section 6.2.3, I thought this approach would best allow the third question, whether or not users' understandings of the maps are modified by the use of sound, to be examined through a close reading of each participant's verbalizations as they worked with each map. Although the other questions could have been addressed through an analysis of the recordings, observation notes, and questionnaires, I thought a finer grained analysis for these questions would also be possible through the cross-sectional analysis.

In addition, as an outcome of the evaluation process conducted by the HOT-Lab researchers, Roberts et al. (2008) provided a summary of usability results for each application and recommendations for modifications and retesting, based on an analysis of the collected audiovisual recordings, the notes, and completed satisfaction questionnaires. The analysis presented here is based on what was in the transcriptions once completed, well after completion of the sessions and my reading of the report by Roberts et al.[14] I provide brief comments comparing results from their analysis with mine at the end of this chapter.

The audiovisual computer screen recordings of the evaluation sessions were transcribed to include the dialogue of the student test participant and the interviewer along with annotations describing the motions of the cursor as the participant worked with each map example. These transcriptions, although time consuming to produce and check, allowed analysis of the sessions to proceed while only relatively rarely resorting to more time consuming examinations of the audiovisual recordings to confirm inferences.

[14] Although I read their analysis when they distributed it, because of my role in organizing and planning those sessions, I was occupied with other work and the transcription of the audiovisual computer screen recordings of the evaluation sessions, in preparation for my own analysis, for much of the year following its release. I did not reread the analysis prepared by Roberts et al. until after having completed my own.

Once the transcriptions were complete, analysis proceeded using computer aided qualitative data analysis software (CAQDAS) to assign codes to sections of text in the transcripts and thereby identify and relate categories of behaviour across the sessions for later comparison, refinement, and documentation. Appendix E contains a table showing codes assigned during the cross-sectional analyses of the evaluation sessions. These code have been used to create arguments documenting the reactions of the test participants to the audiovisual map examples, whether or not they were able to understand and work with the examples, and what difficulties, if any, they had with the examples.

6.3 Analysis of Evaluation Session Data

This section discusses observations concerning the evaluation sessions that highlight either difficulties with using the audiovisual designs or address in some way user interpretations of the maps. Table 6.2 provides a summary of selected analysis discussion points, each designated as 'ad.n,' developed through analysis of the user evaluation session transcripts. For brevity, only a summary of the analysis is discussed in this section. For detailed descriptions of each of the findings listed in Table 6.2, see Appendix E for a list of cross-sectional analysis codes generated during the transcript analysis and details of each analysis discussion point. For reference, Appendix B presents a functional overview of each of the maps evaluated. Throughout the discussion, specific participants are designated as 'Pn' where n shows a participant's position in the chronological ordering of the evaluation sessions.

Table 6.2 is structured using a set of main discussion points each of which generally organized a set of codes assigned during the transcript analysis, but some-

Table 6.2: User evaluations: Summary of analysis discussion points.

Label/Page	Analysis Discussion Point
Ottawa Area Federal Election Sound Map: June 28, 2004 (OAFESM):	
ad.1/p.323	4 of 6 participants decided that the mixing of speeches was significant.
ad.2/p.323	All participants commented on sampled speech content and 3 of 6 disliked that they did not know how to isolate each speech.
ad.3/p.324	Participants used both visual and acoustic information to interpret the sound design.
ad.4/p.325	3 of 6 participants successfully compared district results using sound levels.
ad.5/p.327	Only 1 of 6 participants used information from a recorded speech to answer a question for which only that speech contained a particular perspective on the subject.
Canada's Trade with World Regions: 1976–2000 (CTWR):	
ad.6/p.328	Only 2 of 6 participants decided without assistance that the music varied based on trade data values.
ad.7/p.329	5 of 6 participants were able to understand the sound design using dynamic audiovisual legends if they inferred a relationship between the visual meters and music.
ad.8/p.330	Participants used the visual design when learning the sound design, and 2 of 6 participants continued to use visuals when interpreting mapped data.
Antarctic Territorial Claims (ATC):	
ad.9/p.331	4 of 6 participants disliked that they did not know how to isolate each speech.
ad.10/p.332	3 of 6 participants, when asked, stated that the recorded speeches indicated the overlapping national claims to territory.
ad.11/p.332	The mixed narrations design reinforced contested claims but did not independently signify the concept.
ad.12/p.335	4 of 6 participants listened to and commented on the languages used in the recorded speeches.
Antarctic Exploration (AE):	
ad.13/p.335	5 of 6 participants detected qualitative changes in sound effects without being prompted.
ad.14/p.337	5 of 6 participants recognized the fly-over sound effect as an airplane.
ad.15/p.337	5 of 6 participants interpreted the NASA sound effect as 'space sounds' or RADAR; one had to use visuals to reach this conclusion.
ad.16/p.337	0 of 6 participants related the map background transformation to the "Island of Utopia" narration to which they listened.
ad.17/p.338	Participants interpreted information using both sound and visuals.
ad.18/p.339	The transcripts show two (2) participants interpreting information using sound without acknowledging that source of information.
General:	
ad.19/p.341	Some participants expressed surprise upon hearing sound.

times only one code, into a description of participant behaviour regarding some aspect of the sound design in one of the maps or regarding an aspect observed across multiple maps. For example, in the analysis of *Ottawa Area Federal Election Sound Map: June 28, 2004* (OAFESM), the discussion of participants attending closely to sampled speech content (ad.2 on p.323) is actually a combination of codes highlighting that participants were: a) attending to speech content; b) uncomfortable with overlapped speech, which involved them first trying to listen and then, in some manner, expressing discomfort at having to listen to blended speech tracks; and c) expressing disapproval, with overlapped speech being one reason some participants did so. Codes assigned were not intended to be mutually exclusive and it was possible that the verbalizations of participants or discussions between interviewer and participant could thematically touch on multiple concepts for each of which a segment of transcript would later be coded.

The overlap of codes reflects the different emphases on how sound could be or was intended to be interpreted in each of the examples. For example, the code "overlapped speech: symbolic interpretation" (Table E.1) was originally allocated to tag transcript segments in which participants were discussing the use of overlapped speeches in the Territorial Claims section of the *Cybercartographic Atlas of Antarctica* (ATC) to reinforce the concept of territorial contestation. However, the same code was later also used when it became clear that some participants were interpreting the overlapped speeches of OAFESM as being recordings of House of Commons debates, and in some cases missing the intended quantitative representation of the election results by the gain setting of each speech in that example. During the analysis process, codes are allocated as concepts are inferred from the transcripts but they may later have been modified or merged if further study of the transcripts seemed to warrant it, possibly because very similar concepts appear in other parts of the transcripts. Therefore, in writing the analysis of the transcripts, I have sometimes treated emergent codes as a group to demonstrate a related set of ideas concerning one or several of the map examples.

Finally, the analysis codes and the analysis resulting from their selection were not intended to identify, highlight, or assess *only* usability *problems*. The analysis was also intended, in consideration of this thesis as part of the context for these tests, to allow discussion of participant behaviour in using the audiovisual map examples such as when user interpretations are clearly being based on the visual design, acoustic design, or both. Whereas it may seem intuitive that people will attribute meaning to combined visual and auditory elements of an audiovisual map, the details of interactions between those elements and a user's interactions with both may create meanings in complex ways. Furthermore, cartographic research literature showing empirical evidence of map users' interpretations being based on combined audiovisual design is still very sparse.

6.3.1 Summary of User Evaluation Findings

Without reiterating everything in Table 6.2, this section discusses noteworthy observations about participants' responses to the map examples. A summary of results from these user evaluations are discussed below, organized according to the following topics:

S.1 Sound changes map use because simultaneous acoustic and visual events are interpreted as semantically related.

S.2 Participants were not familiar with audiovisual maps.

S.3 Participants explicitly expressed approval and disapproval concerning the audiovisual map examples, with comparable numbers of occurrences.

S.4 Participants' initial responses to map examples indicate that explanatory materials and user training may be required in addition to complementary visual and sound design.

S.5 By map example, certain notable usability problems require consideration.

Functional overviews of each of the four map applications evaluated are provided in Appendix B. Throughout this discussion they are referred to using the following initialisms: *OAFESM* for *Ottawa Area Federal Election Sound Map: June 28, 2004*, *CTWR* for *Canada's Trade with World Regions: 1976–2000*, *ATC* for the Territorial Claims section of the *Cybercartographic Atlas of Antarctica*, and *AE* for the Exploration section of the *Cybercartographic Atlas of Antarctica*.

S.1 Sound changes map use because simultaneous acoustic and visual events are interpreted as semantically related. The user evaluations of these example maps demonstrated that when participants heard speech and sound effects that occurred simultaneously with visual updates, they interpreted the sounds and visual updates as semantically related. Participants were less likely to consider the music used in CTWR as associated with spatial data than when working with the other sound designs, unless prompted to consider such a possibility. They noticed that the music was changing in conjunction with visual updates as they worked with the map and dynamic audiovisual legends (ad.6 on p.328) but some were unsure what the acoustic changes meant.

The tendency of participants to relate sights and sounds when working with the map atlas examples is apparent in their moment-to-moment interactions: interpreting the sound effects in AE as related to the route maps and expedition leader name associated with the current position of the cursor; interpreting adjusted speech volume levels in OAFESM resulting from moving the cursor between electoral districts as being related to the newly selected district, whether or not they fully understood the encoding of the electoral results in the gain settings; and associating the narrated news stories of ATC with the highlighted country-name hyperlinks or map sectors.

In addition to the above, two examples demonstrate instances during the evaluation sessions in which participants expressed unexpected interpretations based on the cross-modal, audiovisual information in the applications. First, during the evaluation of ATC, P3 was asked to compare the representations of New Zealand and Chile within the atlas. Immediately after being asked this question and as P3 began to move the cursor across the map to find those nations' claimed sectors, all narrations became audible because the cursor moved over the sector claimed by all of Argentina, Chile, and the United Kingdom. Within that mix of narrations, that representing the British claim was coincidentally playing the phrase "Chile has confirmed [. . .]." P3 heard the phrase, interpreted it to indicate that she must have positioned the cursor over the sector claimed by Chile, and asked the interviewer if this was the case. This exchange between P3 and the interviewer is remarkable for two reasons: a) P3 interpreted the explicit reference to Chile as indicating that the cursor was over that country's claimed sector, indicating that she was interpreting the reference as being triggered by her movement of the cursor;[15] and b) although P3 had positioned the cursor over one of the map sectors claimed by Chile, this was signified in the design of the application by the audibility of the Spanish narration rather than the audibility of the British narration or any geographical references within it. P3 incorrectly related the English phrase and her cursor movements as indicating that she had highlighted the sector claimed by

[15]Each of the narrations associated with a national claim to Antarctic territory is played continuously in a loop. Therefore, *what* was being said in the speech as she moved the cursor over a map sector and caused the associated narration to become audible was coincidental.

Chile but she clearly demonstrated that she was interpreting the narrations and visual updates as a combination.

Second, while evaluating AE, P1 started to move the cursor over the expedition selectors array causing the sounds to play and the expedition routes to be displayed on the map. When P1 moved the cursor over the selector for Robert Scott's expedition, the display was updated to show sea routes that, at the scale represented, appear to come ashore and navigate along the coastline, being markedly closer to the coastline than any previously displayed routes in the expedition chronology. In addition, the route display for Scott also shows the first overland routes seen by P1 while scanning the expeditions. Upon seeing this route map, P1 stated "Well I think they're getting closer because at first it's really loud, it sounds like giant waves and it's really loud, but now it's more like it's maybe coast." Because P1 was moving the cursor left to right over the expedition selectors and all expeditions preceding her comment are associated with the seafaring exploration audio loop, the sound hadn't changed markedly although that audio loop does contain some recurring acoustic events throughout that seem loud in comparison to the rest. P1 had concentrated on the visual and textual components of the atlas section while initially examining it and only began discussing the sounds at all once prompted to by the interviewer just prior to re-examining the expedition chronology and making the above comment (ad.13 on p.335). Therefore, although it is possible that P1 was merely reacting to audible changes in the sound loop, her earlier visual concentration coupled with the markedly more noticeable visual route differences accompanying her comments suggest that she said this, at least in part, because of the visual changes.

The tendency by participants to jointly interpret the visual map updates and concurrently played sounds would seem to indicate that Chion's (1994) concept concerning the *added value* arising out of combined seeing and listening among film goers may be applied, albeit cautiously, to audiovisual cartography. Would P1 have heard the sounds getting louder as if the expedition was nearing the coast had she not seen the route map indicating the same? P3 may not have decided she had positioned the cursor over the sector claimed by Chile, or she would have had to base the decision on other information, had she not listened to the British narration. But Chion's concept must be applied cautiously to these examples because his ideas were developed for the linear narrative structure of film. When P3 interpreted what the British narrator was saying simultaneously with her cursor motions as relevant to identifying the Chilean claim, her interpretation of the audiovisual event resulted from a combination of a determinate visual update and an underdetermined sound update due to the effect of timing. If the British narrator had been saying something else, P3's interpretation would have had to be different. Further study of design issues arising from the combination of linear sound segments used in conjunction with non-linear, indeterminate user actions is required.

Finally, participants sometimes interpreted the audiovisual design without acknowledging that their understanding had been influenced by the sounds they were hearing, whereas the explanation that they were responding only to the visual information, as given, seems hard to accept (ad.18 on p.339 discusses two examples of this from the evaluation session transcripts).

S.2 Participants were not familiar with audiovisual maps. Participants were not familiar with audiovisual maps when they began their evaluation sessions. Their lack of familiarity was evident in their surprised responses to the initial generation of sound from the map applications as they began to work with them (ad.19 on p.341) and in their propensity to disregard the sound design of CTWR as potentially conveying quantitative data until prompted to consider the possibility by the interviewer (ad.6 on p.328).

A separate issue from participants' lack of familiarity with the use of sound as part of maps in general, at least potentially, is their lack of familiarity with sounds being used as designed in some of these example applications. In addition to the tendency of participants to not consider the music in CTWR as quantitatively representing trade data,[16] some participants had trouble understanding the significance of how speech was being used in other map examples. Not all participants understood the use of electoral speech volume control in OAFESM as a signifier for voting results.[17] Not all participants understood the significance of overlapped speech in ATC as intended to reinforce the concept of territorial contestation in Antarctica.[18]

By contrast, participants examining the expedition chronology of AE all interpreted the changing sound effects as related to what the expedition was doing while exploring Antarctica, although P1 needed prompting from the interviewer before her attention was drawn to the sounds (ad.13 on p.335).

The greater propensity of participants to understand the sound effects of AE as indicating how the expedition was travelling as compared to the divergent assumptions of participants concerning the purpose of the music in CTWR may indicate the influence of learned listening habits from other media. Although some participants understood that the music signified quantitative measures of Canada's trade, more than half initially stated they thought the music was being

[16]Only two of six (2/6) participants did so without being prompted by the interviewer to consider the sound in relation to the map (ad.6 on p.328).

[17]Four of six (4/6) participants interpreted the gain level used to mix each speech as indicating vote results (ad.1 on p.323).

[18]Three of six (3/6) participants, when asked, stated that the recorded speeches indicated this (ad.10 on p.332).

used for other purposes (ad.6 on p.328), with suggestions including: a) the music was intended to relate a cultural characteristic of the current region, at least until participants realized the same music was used for all regions; and b) the music was intended as soothing background music and not related to the data.

Similarly, the use of the *presence* or *volume* of speech as a signifier of either qualitative or quantitative information, other than that conveyed by an individual speech act, was not familiar to participants. This is apparent in the concerns expressed by participants using ATC and OAFESM about all the narrations being heard at once or about not being able to isolate a specific narration, if they did not discover how to do so (ad.1 on p.323, ad.2 on p.323, and ad.9 on p.331).

Although layering and juxtaposition of different spatial data and the expression of those data through separate sound loops understood as layers may, as it did during the creation of these map examples, seem like a straightforward analogy for visual cartographic layering, and some participants understood each of these examples as such, these applications clearly created listening situations with which some participants were unfamiliar. It would be premature to declare that music may not be appropriate for use in abstractly communicating spatial data or that map designers should avoid allowing multiple narrations to play simultaneously but designers considering such uses must recognize the relative paucity of these kinds of sound designs in popular media and plan to explain the usage clearly for users.

S.3 Participants explicitly expressed approval and disapproval concerning the audiovisual map examples, with comparable numbers of occurrences. While working with the map examples, participants expressed approval or disapproval concerning some characteristic of each of the map examples as shown in Table 6.3. For each map example, the expressions of approval shown indicate statements

Table 6.3: User evaluations: Explicit expressions of approval or disapproval, by map example: '+' and '-' indicate approval or disapproval expressed during participant's map session. Both approval and disapproval in the same session was possible. Greyed out cells show cases where a participant did not test a map example.

	Participant Approvals / Disapprovals								Totals	
	P1	P2	P3	P4	P5	P6	P7	P8	Approvals	Disapprovals
Election	+ -		+	-	-				2	3
Trade	+						-	+	2	1
Claims		-		-	+ -		+		2	3
Exploration	+ -				+				2	1

made by participants that they liked the use of sound, possibly in conjunction with statements that they liked the visual design or the interactive behaviour of a map example. Some approvals were qualified. P5 stated that the narrations "enriched" the information concerning national claims to Antarctic territory but suggested that he would like to have English versions of each speech included and expressed dislike for the overlapping speeches. P1 initially approved of the sound playback initiated by the Antarctic expedition selectors in conjunction with the visual map updates, declaring "that's so cool!" But prior to discovering any sound other than that associated with the ocean voyages, she fairly quickly suggested that the sound of waves was becoming tiresome (ad.13 on p.335).

Expressions of approval were made concerning all maps equally as shown in the totals of Table 6.3. These expressions of approval are difficult to interpret, possibly being reactions by participants to any aspect of the design (visual, acoustic, or interactive behaviour) or its novelty, or to the information conveyed. Because the participants' statements coded as *approval* or *disapproval* during analysis all concerned the use of sound, it may be possible to interpret them as indicating a general acceptance for the use of sound in audiovisual maps among this group of participants. Both because of sample size and because the participants self-selected to participate in evaluations of "maps using sound" (Appendix D), it is impossible to suggest that this openness would extend beyond this group of participants.

The expressions of disapproval shown in Table 6.3 were all made with respect to specific characteristics of the examples evaluated by the test participants. The participants expressed disapproval most often for the uses of overlapped speech in ATC and OAFESM, with three participants expressing disapproval concerning this in each of these two maps (see ad.1 on p.323 and ad.9 on p.331). For example, when first positioning the cursor over the contested map territories in ATC, P4 complained that "all these people came up at once and started talking" (ad.2 on p.323).[19] This may indicate that, for some users, either too much information was being presented at once and cognitive load was becoming an issue or that overlapped speech was simply unusual in their listening experience.

In addition, three statements of disapproval identified other concerns with an aspect of one of the map sound designs. First, P7 initially stated that the music of CTWR was distracting prior to the interviewer encouraging him to consider the sound in relation to the range of data classes using the dynamic audiovisual legends. He subsequently interpreted trade data successfully using the sound design of the application but still said that he thought the sounds were "a little bit distracting," indicating both that his disapproval remained and that cognitive overload could have been a problem for P7 while using CTWR.[20]

Second, P1, having not understood the quantitative control of electoral speech sound levels, said that "it really bothered her" that she heard the speech of the Conservative Party leader when her cursor was positioned over a district coloured to indicate that the Liberals had won the election (ad.2 on p.323). Rather than a

[19]P4 voiced a similar concern when first hearing the combined election speeches of OAFESM, and using similar language (ad.9 on p.331).

[20]As already discussed (ad.6 on p.328), participants working with CTWR frequently needed assistance or prompting, first, to consider that the sound design may signify spatial data and, second, to understand the design and related trade concepts.

comment that necessarily expresses dissatisfaction with the sound design of the map, this comment can be interpreted as indicating that P1 did not understand the overall design of the map and needed a better explanation of the intent.

Third, P1's concerns about the sound associated with seafaring expeditions in AE becoming tiresome were already mentioned above. But this participant's verbalizations while working with the map focused predominantly on its visual components and their updates resulting from her interactions (ad.13 on p.335). Remarkably, once she began to initiate playback of sounds other than the wind and waves loop associated with the seafaring expeditions, she did not comment on them or seem to notice them until prompted by the interviewer to consider the sound design. This may indicate that this participant prefers visual information, or her initial failure to perceive and comment on the auditory changes could have resulted from cognitive overload (induced by the number of interactive updates within the application in combination with her own *think aloud* verbalizations). Alternatively, having already decided and stated that the sounds were tiresome, she may simply have stopped attending to them.

The generalized approval, possibly interpretable as an openness to the use of audiovisual maps, combined with the specific statements of disapproval, each identifying a design flaw from the perspective of that participant, suggests that audiovisual maps could have a role but problems with their design and use must be resolved.

S.4 Participants' initial responses to map examples indicate that explanatory materials and user training may be required in addition to complementary visual and sound design. Considering the paucity of sound use in cartographic designs, when it is used it must be explained. CTWR presents spatial trade data using audio indicators with no explanation. AE similarly contained no description of what the sounds associated with the expedition selector array were intended to indicate. The textual description accompanying the map in OAFESM stated that the speeches were "adjusted according to the party's" election results without saying what adjustment was done. By contrast, the textual description in ATC stated that the sound was intended to "highlight the tension" without explaining how.

Four of six (4/6) participants working with CTWR did not consider the music as being related to spatial variables until prompted to do so, even if they heard changes in the music while working with either the map or the dynamic audiovisual legends (ad.6 on p.328). Although five of six (5/6) participants indicated that they heard changes, only two of those five determined that the music was being controlled based on the trade data without the interviewer prompting them to consider this. The other three of the five (3/5) initially seemed to focus on the visual meter updates and continued to look to those while they worked with the map until an interviewer question or comment caused them to consider the significance of the music. Once participants were prompted to consider *that the sounds were related* to the spatial data, working with the legends, possibly while discussing the sound changes with the interviewer, was effective for all but one participant, in helping them *to understand how* the sound changes related to the data.[21] Because some participants needed to be prompted to even consider a relationship between the audio and the data, and others had difficulty determining how the music varied with the data, the dynamic audiovisual legends, alone and without explanation, are insufficient for some users to understand the sound design.

Although AE contained no explanation for the sound effects used, the qualitative, descriptive selection of sounds was interpreted by the participants as sig-

[21] P6 was unable to verbalize how the instruments for the trade balance indicator were being controlled in relation to the data (ad.7 on p.329).

nifying an aspect of the explorers' experiences (ad.13 on p.335). In response to interviewer questions about the example (e.g., who first trekked overland?), participants responded in a variety of ways, some relying on listening, some on the visual map, some by reading, and often by using information gained or remembered from one modality and subsequently confirmed through another (ad.17 on p.338).

Although the sound design of AE was intended to reinforce the visual and multimedia design more than to necessarily add new information (Caquard et al. 2008), that objective was only achieved when participants understood what the sound represented, possibly because of supporting visuals. But difficulties in recognizing and inferring the meaning of sounds was apparent for some participants (ad.14 on p.337; ad.15 on p.337). This was especially notable with the (imaginary) sound associated with NASA's imaging of the continent. P4 used the text panel or visual icons on the map to clarify this sound after initially suggesting that it was a submarine. P1 stated that she did not know what the sound represented.

When asked to compare election results between electoral districts using OAFESM (ad.4 on p.325), two of six (2/6) participants initially attempted to answer by interpreting the relative sizes of electoral districts on the visual map as indicating the vote percentage of the winning party in each district. This is a misinterpretation of the categorical map but may have been a guess made by the participants when they otherwise thought they had not been given information, or usable information, with which they could answer the question.[22] There is no indication in the data collected for these users' sessions that either read the text

[22] These two participants were not given assistance in understanding the audiovisual design of the map, despite the transcript showing that they both noticed the relative volumes of the leaders' speeches being adjusted without being able to explain, when asked, why more than one voice was audible. Instead, they each commented that this made the voices difficult to listen to or "confusing" over certain regions.

describing the use of sound in the example.

Finally, three of six (3/6) participants working with ATC did not understand the overlapped sound design as qualitatively signifying the competing national claims to territory (ad.10 on p.332), despite the text panel for the atlas section containing a seemingly clear textual description of the purpose of the narrations.

Thus it seems that careful design and possibly training is needed to ensure that as many users as possible understand the purpose behind an audiovisual map design. The dynamic legends alone, although visual and interactive, may not be sufficient for all users to consider and understand quantitative uses of sound. Text may not be read. Combining the dynamic legends with text may then be necessary but risks simply creating a larger block of explanatory material that at least some users will ignore.

During the map evaluation sessions for CTWR, the interviewer often realized that a participant was not attending to or could not explain the sound design. Quite often, this was followed by a brief interchange during which the interviewer encouraged the participant to initiate the sounding of the auditory representation of different variable classifications using the dynamic audiovisual legends and to describe the changes they heard. Once this was done, five of six (5/6) participants seemed quite comfortable working with the sound design in relation to the map, with and without the dynamic audiovisual legends being displayed. This highlights that training could be quite effective in drawing user attention to the use of sound designs in audiovisual mapping, in addition to the use of design approaches that incorporate textual explanations and other learning aids.

The descriptive approach to sound design used in AE, generally conceived as a simple, non-detailed cinematic approach retained from an earlier animated version of the exploration time line,[23] and the participants' relative ease in understanding the significance of that design suggest that familiarity from other media with descriptive sound designs may lessen the explanatory requirements in such cases.

S.5 By map example, certain notable usability problems require consideration. Table 6.4 lists usability problems identified by this analysis of the session transcripts, by map example, along with suggestions to remedy or alleviate each problem. These problems and suggested solutions would be candidates for consideration during iterative development of each of these map examples. Because further development of these map examples was not going to happen, these problems have been listed here and were taken into account in the design of a new audiovisual map application, related thematically to CTWR, to be discussed in Chapter 8.

6.4 Discussion of User Evaluations

The user evaluations were designed to address the questions outlined in Section 6.2.3. The above summary of analysis concerning the evaluation session transcripts, along with the detailed discussions in Appendix E, has addressed those questions and has produced a useful summary of problems and insights into how these participants responded to the map examples. Considering the sample size used, these results can't be generalized to other audiovisual map examples but can be used to generate hypotheses for further testing. The application of any concerns and insights raised by these evaluations in new projects will involve project-specific design and careful consideration of project objectives.

Out of these evaluations, new ideas for audiovisual map sound designs have

[23] See Section B.3.2 and analysis discussion ad.13 on p.335.

Table 6.4: User evaluations: Key usability problems, by map example with number of occurrences for each. Continuing on next page ...

Occurrences	Problem / Suggestion for Resolution
Ottawa Area Federal Election Sound Map: June 28, 2004 (OAFESM):	
3/6	**Participants disliked that they could not isolate (disentangle) the mixed speeches**: The legend designed as part of the map provided the capability of listening to each speech separately but this was not explained in the design of the map. Three participants did not discover this capability and expressed frustration with the mixed speeches or their perceived need to concentrate to listen to one without assistance from the interface. This problem could be addressed through training or a clear explanation of the function of the legends; dynamic audiovisual legends, developed after this example, could alleviate this problem.
2/6	**Participants did not, without assistance, understand the quantitative meaning of the speech gain levels**: The text accompanying the map example stated that the audio was adjusted according to election results as a user moved the cursor over the map but did not say how it was adjusted. Clarifying the text may help some users (if they read it); using dynamic audiovisual legends may improve this because this design provides an opportunity to visually show the election results in relation to the varying sound design; user training could help.
3/6	**Participants did not successfully compare district electoral results using the sound design**: One participant never considered using the sound design to compare the election results between districts, using the visual map design instead. Two participants that did attempt to use the sound design to compare results were unable to answer. Training and the use of dynamic audiovisual legends could alleviate this problem.
1/6	**Participant interpreted a negative statement about the Liberal Party in the Conservative leader's speech, audible while the cursor was positioned over a map district coloured red to indicate that the Liberals had won the election, as contradicting the visual map**: This participant perceived this to be an error in the design of the map. This is a misinterpretation of the sound design for the map. This should probably be alleviated by attempting to clarify the intention of the sound design for users. If the design intent is understood, the inclusion of partisan comments by politicians within the selection of narrations used in the sound design is appropriate to the map theme and assists in identifying the party affiliation of the speaker.

Table 6.4: User evaluations: Key usability problems, by map example with number of occurrences for each (continued)

Occurrences	Problem / Suggestion for Resolution
Canada's Trade with World Regions: 1976–2000 (CTWR):	
4/6	**Participants did not understand the relationship between trade data and the control of music while working with the map and dynamic audiovisual legends until prompted by interviewers**: The application interface contained no explanation of the use of sound but the initial view of the application shown to participants included the dynamic audiovisual legends. This alone seems to have been inadequate to help some participants understand the intended use of sound design. Explanatory text may help; training videos, available through help hyperlinks from the map page, explaining the use of the legends map alleviate the problem.
2/6	**Participants preferred to work with the dynamic audiovisual legends displayed**: The dynamic audiovisual legends in this map example were designed to be placed in a "learning" view of the application a) to save screen space for the map; and b) to allow evaluation testing to see if participants could interpret quantitative data using sounds. Users could be allowed to use both sound and visuals by redesigning the dynamic audiovisual legends to be smaller and including them on the main map page. This was done in the later map example *Canada-USA Commodity Trade 1976–2000* (see Chapter 8).

Table 6.4: User evaluations: Key usability problems, by map example with number of occurrences for each (continued)

Occurrences	Problem / Suggestion for Resolution
Antarctic Territorial Claims (ATC):	
4/6	**Participants disliked that national claims speeches played as a mix and they didn't know how to isolate (disentangle) each speech**: Hyperlinks in the text panel initiated playback of each speech on its own but this was not explained in the interface. Clarifying this in the text could help but the side panel already contains a fair amount of text so this will not solve the problem for all users. Introducing a graphic legend on the map itself that allows each speech to be played alone may help but a simpler design than the dynamic audiovisual legends used with CTWR may be acceptable because the speeches in this map do not have a quantitative meaning. Such a legend could also offer a transcription and, in the case of the Spanish speeches, a translation of the narrations which were also suggested by some participants.
3/6	**Participants did not understand the use of speeches to reinforce the concept of overlapping national claims to territory**: The text panel accompanying the map contained a statement that the sound design was intended to reinforce the contestation of territory but did not say how this was done. It is possible that if all perspectives on the territorial claims were outlined, for example through the translation of the Spanish narrations, the intention would be clearer. Considering that the map design was intended to use sound in a slightly unusual way to capture user attention concerning the contestation of territory, it is conceivable that this intention would be undermined by over-explaining it (like over-explaining a joke).
1/6	**Participant interpreted a statement in a recorded speech as being the result of cursor movements over the map rather than just the audibility of the speech containing that statement as resulting from the cursor positioning**: One participant, P3, heard the British narrator mention Chile by name and assumed that this indicated that her cursor was over the section of Antarctica claimed by Chile. This sound design, based on a set of looping narrations being muted or made audible based on logic driven by a user's interactions with the map, if not understood as in this case, could cause unpredictable interpretations by users. This problem could be alleviated by designing the interface to more clearly communicate the concept of territorial contestation, possibly by making transcriptions and translations of all narrations available as suggested above.

Table 6.4: User evaluations: Key usability problems, by map example with number of occurrences for each (continued)

Occurrences	Problem / Suggestion for Resolution
Antarctic Exploration (AE):	
2/6	**Participants had difficulty interpreting the sound associated with NASA's imaging of Antarctica**: This sound effect is a fabrication intended to convey the idea of a satellite orbiting and collecting earth images. As such, the lack of certainty expressed by participants concerning the meaning of the sound highlights the potential difficulties of this design challenge in some cases. This problem could be alleviate in this map example by designing the interface so that positioning the cursor on the expedition selectors also shows the transportation icons. Currently, these icons only appear when a user clicks on the expedition selectors.
2/6	**Participants were momentarily confused by the "dead air" at the beginning of the fly-over expedition sound effect before the plane became audible**: Audio loops designed to be made audible under certain logical conditions should be edited to remove "dead air" when possible to ensure the user hears the sound when their actions trigger playback.
1/6	**Participant identified the plane sound effect as being the sound of a helicopter**: This is not a major concern since both sounds indicate aerial surveillance of the continent. It does highlight the tricky problem of designing and using sound effects so that they are understood. See the discussion of the NASA sound effect above for a suggestion to alleviate this problem.

been generated and applied in the design of a more recent application, related thematically to CTWR, for which Chapter 8 provides a brief overview. In addition, design criteria and objectives for the cartographic sound subsystem used in the creation of audiovisual maps have been reviewed and updated as a result of lessons learned during and in preparation for the evaluations and will be discussed in the context of a general description of that subsystem in Chapter 7.

The number of participants in these evaluations was small but sufficient to provide reasonable task coverage for usability testing of each of the evaluated audiovisual map examples, each of which supports only a small range of possible user tasks in which sound could play a significant role. OAFESM was designed as a small standalone demonstration of the possibility of using election speeches as the basis of a representation of electoral results, using a fixed summary of electoral data, and offers relatively few tasks or data points from which to choose. CTWR uses data that is highly aggregated, spatially and temporally, and the version evaluated, as discussed in Appendix E, was further simplified, thereby also reducing the representable data set and the range of possible test tasks. Similarly, the two examples from the *Cybercartographic Atlas of Antarctica*, although more complex and complete map examples, each used sounds in very specific ways that again limited the possible range of tasks relevant to user evaluations of sound use.

The objectives for these evaluations stated that understanding the *usefulness* of the sound designs in these audiovisual map examples was also important, in addition to identifying usability problems. The analysis summarized here highlighted ways in which the sound designs can be useful, including: 1) *influencing interpretation and/or reinforcing visual information*, 2) *signifying additional information*, and 3) *engaging users*. First the evaluations showed that sounds influenced interpretation, for example when P8 described the exploration sea routes as "stormy" (ad.18 on p.339), and reinforced visual designs, for example, when P7 described Antarctic territorial contestation as "heated" after hearing the narrations (ad.12 on p.335).

Second, the evaluations showed that users can discern quantitative information based on acoustic design, allowing the possibility of offloading users' visual perception, as was done with music in CTWR and speeches in OAFESM. But these evaluations also demonstrated how design challenges can limit the ability of users to benefit from information and highlighted that training, in addition, to good design may be necessary.

Third, the evaluations showed that at least some participants enjoyed listening to the sound designs during their initial encounter, based on their statements of approval. But personal tastes vary and components of the sound design that were enjoyable for one participant may have been unpleasant for another. For example, the music in CTWR was assessed differently by P7, who found it annoyingly "happy" and distracting, and P8 who stated that she liked listening to it (at least for the duration of the evaluation). Allowing users to select from a range of musical options or indeed supply their own music, at least in sound designs such as CTWR in which the sound is being used as an abstract signifier, may be useful.[24]

It would be premature to identify any particular use of sound or style of sound design as a best practice or a particularly bad practice, although the analysis of these evaluations has highlighted the importance of considering the expectations of users along with the necessity of careful design and the advisability of training, when possible, to prepare users to effectively work with audiovisual maps. Some participants had difficulty with aspects of the maps evaluated but sound design

[24]In fact, the full version of CTWR does exactly that, allowing users to select from 8 possible sound designs, some of which use different musical compositions. This feature was deemed an unnecessary complexity when scripting the user evaluations and the feature was hidden during the sessions.

ideas that caused problems in these maps may be valid in other maps, if better explained, if training is provided prior to use, or if the design presents blending sounds to users gradually or in a manner that clearly provides them with control of the experience. This is not to say that there were no lessons to be learned from these evaluations. Overlapping speeches, for example, clearly must be used carefully giving consideration to what appears to be a common expectation among users that they be given obvious means with which to isolate and hear each speech, and understanding that there are predominant models of sound design from other media that influence user expectations. But in designing audiovisual maps for specific purposes, as with other communicative media, it is difficult to say that there will not be opportunities in which designs that run counter to users' expectations would create a memorable experience or an effective learning situation. If, or when, audiovisual maps become more common, users expectations are quite likely to shift as they become familiar with common practice as it develops, perhaps creating a shifting cultural, economic, and technological context in which user expectations and map making practice develop in tandem.

Understanding the usefulness of audiovisual designs as they relate to differing characteristics of the population, such as gender or age, would require a test design that purposefully selected participants from across such groups and would probably require a combination of methods to include both qualitative and quantitative analysis of user response and capability with the maps.[25]

Although usability testing with analysis designed to determine compliance

[25]Abilities for spatial perception and cognitive manipulation of spatial representations are widely recognized as subject to gender differences, although the sources of the differences (e.g., inherent abilities, attitudinal differences, positive reinforcement, the availability of role models) are the subject of continuing study and controversy (Quaiser-Pohl and Lehmann 2002). The use of some types of auditory material (e.g., narrations) like the inclusion of multimedia in maps could reduce the emphasis on spatial abilities in some types of map use, for example, in comparison to navigational and orientation tasks with the aid of a map.

against criteria, often quantitative, outlining project objectives such as release quality objectives is clearly useful for some projects, the examples evaluated here were not really being developed within such a context or, in the case of the examples taken from the *Cybercartographic Atlas of Antarctica*, the ongoing project context in which that kind of ongoing development cycle was being used had ceased to exist. Therefore, for this thesis research, my main concern was to focus on the process data generated during the evaluations to attempt to understand the usability and usefulness of the sound designs as exhibited in moment-to-moment interactions of the participants with the examples, determined by their comments and discussions with the interviewers.

The person analyzing the results of a set of user evaluations can impact the type and number of issues identified (Jacobsen et al. 1998). Despite differences in method and focus, the analysis conducted here and that reported by Roberts et al. (2008) are different in detail but consistent in overall tone and assessment. Despite the focus of Roberts et al. on the overall usability of each application compared to the focus of my analysis on usability and usefulness of the sound designs with an understanding that overall usability would impact users' experiences with the map examples and therefore the resulting analysis, both argue that participants were open to the use of sound but had specific problems with details of the application designs. The analysis presented here is more detailed in its examination of participant responses to the sound designs and how those influenced their interactions with the maps. These methods are complementary and future research concerning audiovisual map applications within an iterative UCD process would benefit from the application of both approaches, although selective use of the time-consuming transcription process would reduce costs and provide for more timely feedback from the qualitative analysis.

Research questions relevant to the responses of users of audiovisual maps that could be addressed through user evaluation studies remain and would be useful in addition to the types of questions studied in the testing reported here. Such topics include: 1) *longitudinal studies of audiovisual map use*, and 2) *the effectiveness of training and explanatory materials*.

First, longitudinal studies are required to better understand users' abilities to learn and gain confidence with auditory representation and to assess the requirement for variability in the sound design to avoid the potential for user boredom and disengagement. Such a study requires an application that uses sound in the presentation of subject matter that has sufficient breadth or depth to allow repeated sessions with individual users to be productive.

Second, as suggested by this research, methods are required to ensure that users of audiovisual maps take notice of the relevance of an application's sound design. Testing of the effectiveness of various methods (e.g., text descriptions, training videos, lists of frequently asked questions) are required. As shown in the evaluations of CTWR, guidance from the interviewer was effective in helping participants to attend to the sound design but it may be uneconomical and unrealistic to consider personal interaction as a method for training users concerning a web technology such as is proposed here.

7

A Sound Subsystem for Audiovisual Maps

As part of the research for this thesis, I designed and wrote software to create a sound subsystem for use with web-based audiovisual maps and atlases.[1] This chapter provides a high level overview of the design and use of the sound subsystem, and then discusses the sound design created for *Canada's Trade with World Regions: 1976–2000* (CTWR), one of the map examples used in the user evaluations discussed in Chapter 6, for which Appendix B provides a functional description.

The sound subsystem design as discussed here was completed after the user evaluations discussed in Chapter 6 in recognition of the need for greater flexibility in the specification of sound designs for audiovisual maps. The software as previously structured defined an *auditory indicator*, the interactive acoustic representation of a spatial variable, as a single construct defining the sound(s) to be used and how the parameters controlling the playback of the sound are manipulated in response to user actions while working with the map. By contrast, the design dis-

[1] Early versions of the software were included as part of the *Nunaliit Cybercartographic Atlas Software Framework* (Hayes 2006) developed by Carleton University's Geomatics and Cartographic Research Centre (GCRC). Appendix F provides a brief history of the development of the sound subsystem software implementation and notes the contribution of Jean-Pierre Fiset, a colleague of the author's at the GCRC. Although the analysis of requirements for the sound subsystem and its design, as described here, were the work of the author, J. P. Fiset made important contributions that extended the range of uses of the sound subsystem and, most importantly, improved the responsiveness of the sound subsystem as part of interactive maps. See Appendix F for more details.

cussed here is more flexible and reusable than the earlier software model and better supports an iterative, experimental process of creating and testing sound designs as part of the creation of an audiovisual map. This modular sound subsystem was created while I began work on a new interactive audiovisual map example, to be discussed in Chapter 8, intended to allow me to explore design ideas that emerged in part from lessons learned through the user evaluations discussed in Chapter 6. Although the map examples used in the user evaluations predate this design, I retrofitted the new sound subsystem design into CTWR as a way of testing the new software as I wrote it. The discussion in this chapter then presents the sound design for CTWR in terms of the new sound subsystem design but the functional characteristics of the sound design presented are identical to those discussed in Chapter 6 and Appendix B.

Figure 7.1 depicts possible means by which parameters could be controlled to modify the performance of a prerecorded sound or a predefined synthesized musical composition. The examples represent mapped variables using audio in which: a) scale transposition is used to adjust playback of a synthesized instrument into different registers to represent values of a nominal variable; b) the combined mute settings of a pair of instruments represent different states of a nominal variable; and c) distinguishable sound levels (gain settings) of a recorded sound effect (or other type of audio such as music or voice) represent distinct values of an ordinal variable.[2] The option represented by Figure 7.1b will be discussed in more detail

[2]In comparison to Krygier (1994) who recommended the use of sounds transformed by one audio parameter (e.g., each of timbre or pitch separately) to encode data, the approaches shown in Figure 7.1 could use only one transformation parameter (a: register) or may use multiple (b: timbre, register, duration) by working with the structure of music. According to Levitin (2006, 73–80), many of these sound parameters are used together in the cognitive separation of sounds (e.g., identification of different sources of sound) and in the grouping of sounds (e.g., identifying a group of notes as being created by a single instrument). Using music in this way may then allow users to more easily separate multiple linked sounds if used together in a design. More testing would be required, including comparative tests. With this type of representational encoding, a

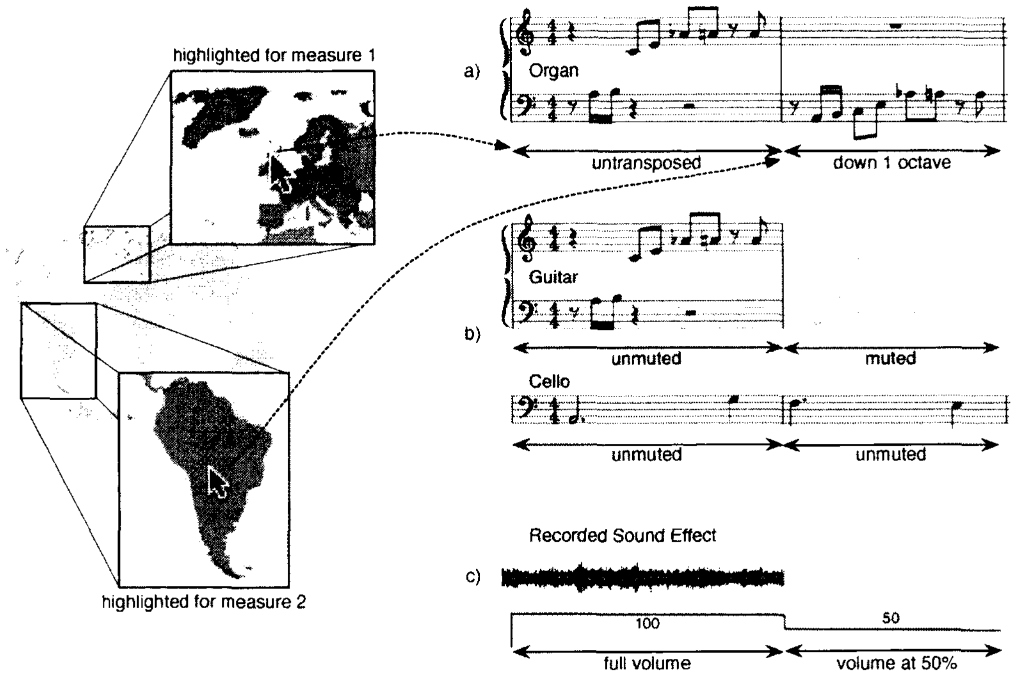


Figure 7.1: Linked audio: Selected design possibilities (figure concept and design: Glenn Brauen).

below because, although each of the depicted audio control methods has been used in at least one of the user selectable sound designs in CTWR, the version evaluated by users and discussed in Appendix B employed this approach to create a classified auditory representation of the *balance of trade* variable. Many of the selectable sound designs in CTWR use variations of the approach represented in Figure 7.1c (i.e., *louder means more*) to represent the *value of exchange* variable.

The implementation of linked audio designs such as those shown in Figure 7.1 require that: 1) *the sounds are modular,* 2) *the sounds are variable,* and 3) *the sound subsystem as a whole is responsive to interactive or scripted control.* First, the sounds must

user may listen to many insignificant audible changes (e.g., each note in a melody) prior to hearing a transition in the audio (e.g., a tempo change or key transposition) that actually signals a variable change and so may have to spend considerable cognitive resources listening. Although some of these linked types of parameter control have been tested and reported (e.g., scale transposition in Krygier 1994), some of these proposals for parameters to be controlled, types of sounds to be used, and implications for interpretation related to particular characteristics of sounds chosen require testing to understand how well they are accepted and understood by users.

be modular elements of a composite overall design with each sound (e.g., a specific recording of narration or a snippet of birdsong) being played or stopped according to the representational design relating that sound with mapped data. Not only is this a precondition for variability in response to user actions, but it also allows sound designs for independent spatial variables to be developed, tested, and evaluated separately. If the design for one auditory indicator seems satisfactory but others are not, the design for the first can be frozen while alternative designs for those requiring improvement are developed and tested.

Second, to signify changing variable values in response to user actions such as cursor movements or as part of a scripted representation of a sequence of map states, each auditory indicator must be able to vary its acoustic performance as required while a user works with the map. The indeterminate, non-linear interaction process of a user working with a map does not allow the possibility of preloading a correctly processed auditory representation. In addition, all of the auditory indicators, and any other components of the sound design such as narrations or themes, must be reproducible as a composite sound design, despite its construction based on modular sound components.

Third, as part of a complete audiovisual map application, the sound subsystem must accept cue requests identifying an auditory indicator and, in some manner, specifying the acoustic parameter changes required of that indicator. Sound designs must be able to react to a user's actions with minimal delay if the user is to interpret modifications in the sound as related to their interactions with the audiovisual map and with the visual map updates likewise initiated by their actions. Although the active gestural involvement of a user moving a cursor over the visible surface of an audiovisual map can be thought of as the active direction of sounds by the user, the temporal sequencing or animation of a sequence of audiovisual map states, requiring the scripting of both the acoustic changes and the visual transformations of the map, requires automation of both modalities.

These characteristics stress that the map as distributed, and the sounds designed as part of it, must retain digital media characteristics as outlined by Manovich (2001, see also Chapter 3). Modularity, variability, and automation are addressed in the discussion above. The desire to distribute and use audiovisual maps over the World Wide Web, an instantiation of digital telecommunication technologies, requires each of the predefined sounds to be numerically encoded. Numerical representation is then the basis of variability in this technology context.

The possibility of sound designs as an effective representational strategy relies on the potential for consistent transcoding of those designs. The computer system to which an audiovisual map is distributed must have the ability to transform each of the numerical sound definitions into a reasonable reproduction of the sound intended by the audiovisual map designer and must be able to control the playback of the sounds according to the variability intended for the representation, thereby allowing the possibility for that digital artefact to be understood as a cultural object. The distribution of the sound design as part of the audiovisual map over the World Wide Web places all of the elements of that design simultaneously both in a cultural context, as sounds intended to be heard and understood as such, and in a technological context, as numerical representations that can be transmitted across a network and are subject to the behaviours of the network and computer systems sending, receiving, and forwarding them.[3]

[3] The transcoding of a predefined sound when transmitted over a network requires more than that the receiving computer be capable of translating and varying the numerical representation as intended. It may also be dependent on the network connecting a map user's computer to the system hosting and distributing those sound definitions being able to transmit those sounds in a timely manner. Sampled audio requires a significant amount of information to accurately reproduce a sound and thus requires a network that is capable of transmitting information quickly and consistently. Sampled audio may be used as a *stream* whereby the sound reproduction begins at a

Extending the design overview presented in Figure 5.1, Figure 7.2 provides an abstract overview of an audio indicator in the sound subsystem. An audio indicator is created using one or more *audio sources*, each a prerecorded digital audio file or an instrument in a synthesized composition, and shown as circles in Figure 7.2. Each audio source for an auditory representation is connected to the computer's audio output through a series of *audio effect processing components*, shown as squares in Figure 7.2, each of which typically manages a single parameter of that source's sound output (e.g., gain, pan, mute, or play/stop settings). The source and effects processing set is controlled by the current value of an associated spatial variable, with a *variable monitor*, shown in Figure 7.2 as a circular sector, passing detected value updates for the variable through a *classifier object*, the output of which is then passed to all of the audio effect processing components to modify the output of the audio indicator.

In Figure 7.2, control updates are designated using dashed lines indicating the run-time reconfiguration of acoustic control or processing parameters (e.g., gain, start, stop, playback sample rate) in response to variable value updates. A variable value update causes the new value to be passed to the classifier, resulting in an updated control parameter being passed to all of the audio effect processing components used as part of the audio indicator. The solid lines in Figure 7.2 show the audio data path running sequentially through the set of effect processing compo-

user's computer before the entire sound definition is transmitted, if indeed it is possible to think of the stream as being a finite sound definition (e.g., Internet radio streams are unbounded). Conversely, sampled audio may be used as a *clip*, whereby it is transmitted and stored on a user's computer before being played, with the understanding that playing the file multiple times only requires that it be transmitted once. Musical Instrument Digital Interface (MIDI) compositions, in comparison to sampled audio, are very compact descriptions of performance but require more of the receiving computer in the form of a functioning MIDI implementation and an installation of the instrument sound banks for which the composition was designed, in addition to the digital-to-analog-sound conversion capabilities upon which sampled audio playback is also based. See Caquard et al. (2008) for a discussion of issues concerning the technological context of sound as part of audiovisual maps on the World Wide Web.

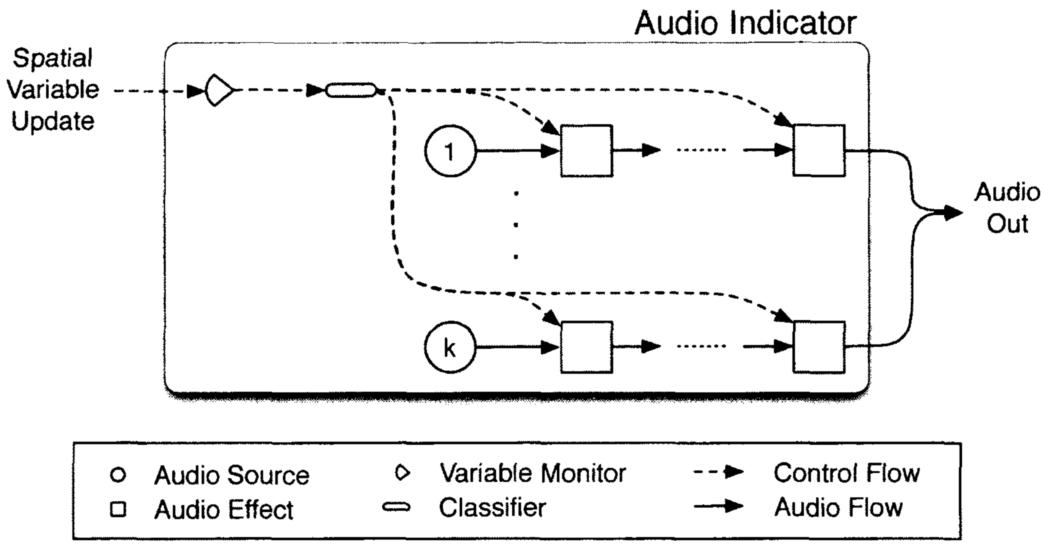


Figure 7.2: Audio indicator components (figure design and production: Glenn Brauen).

nents, the flow of which is broken by an ellipsis to show that the audio indicator design supports a variable number of effects processing components that can be configured as needed.

The sound subsystem currently uses two types of classifiers, each generating one of two types of control parameters: *classified* or *linearly interpolated*. The classification object for a classified control parameter is configured as part of defining an audiovisual map by assigning it a set of value ranges, each range of values associated with a class index ($0 \ldots n-1$, where n is the number of classes). A value update causes the classifier to determine the class of the new value and the corresponding class index is then passed to each of the audio effect processing components, each of which has been correspondingly configured with a set of audio parameter controls, selected by class index, as shown for a simple gain controlled audio indicator in Figure 7.3a. Gain parameter values range from 0.0 (inaudible) to 1.0 (loudest). Therefore, the example shown creates a five class audio indicator in which the loudness of the associated sound increases with the class index. Not

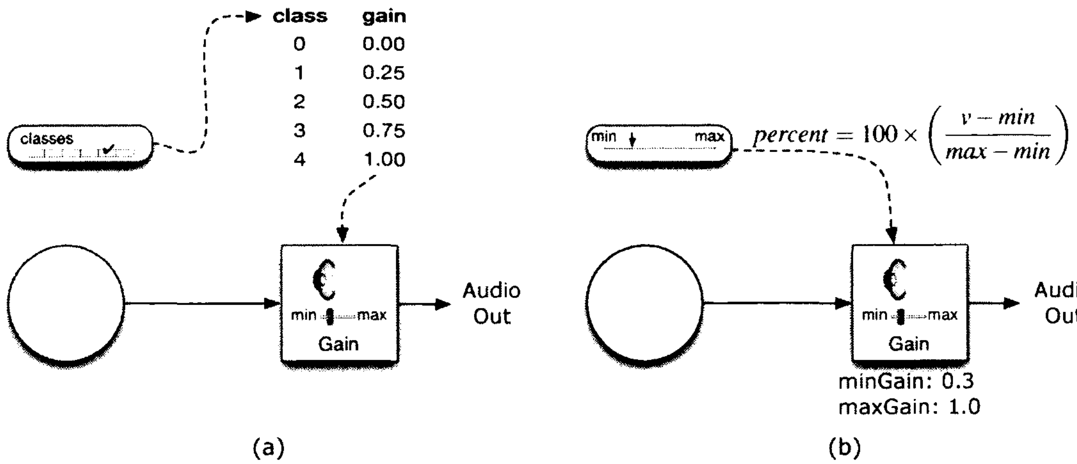


Figure 7.3: Acoustic parameter controls example: a) classified and b) linearly interpolated (figure design and production: Glenn Brauen).

shown, the defined class value ranges for a spatial variable can be configured as required over real or integer number ranges.

The classification object for a linearly interpolated control parameter is instead configured with the minimum and maximum values for the expected data range. When a value update is received, the new value is linearly interpolated to determine its offset from minimum within the configured data range as a percentage, as shown in Figure 7.3b. The percentage value is then passed to each of the audio effect processing components, each of which has been configured with the details necessary to compute a new control parameter value based on the received percentage. As shown in Figure 7.3b, a linear gain control is configured with minimum and maximum gain settings and the percentage is then used to compute gain control values within the range defined by the configured minimum and maximum. In the example shown, a variable is represented by adjusting the loudness of the associated sound with greater data values causing the sound to be played more loudly but the sound is never completely inaudible in this example because a lower limit of 0.3 has been configured for the gain setting.

An audio design for an audiovisual map in the most general case then includes a set of audio indicators, each associated with a different variable (and possibly other sounds as discussed below). The user interface components of the map application provide event handling routines that detect cases such as the cursor being moved over certain regions of the map or dynamic audiovisual legends.[4] These event handlers in turn are written to issue function calls to the sound subsystem application programming interface (API). The sound subsystem API provides interfaces that can be called to update the variable monitor created for an audio indicator to a new value, thereby triggering sound subsystem processing, as depicted in Figure 7.2, leading to an altered sound output for that indicator.

The variable monitors, as shown in Figure 7.4, provide a central point through which all sound subsystem updates related to these variables are dispatched. This allows the sound subsystem to be used as a flexible module, enabling optional components such as the dynamic legends to be added easily or left off if unnecessary to a design. If added, the dynamic legends are simply connected to the variable monitors both as registered listeners (i.e., the dynamic legend asks the variable monitor to notify it of any value updates so that the legend can show visual updates tracking the variable changes) and as an updater of the variable monitor, notifying the monitor when a user directly manipulates the dynamic legend as a controller, as shown in Figure 5.2a. Other user interface controls such as mute buttons or user interface gain adjustments (for the map as a whole or for each auditory indicator) can interact with the sound subsystem design.[5]

---

[4]Dynamic audiovisual legends are discussed in Chapter 5 (see Figure 5.2). See Appendix B.2 for a discussion of the dynamic audiovisual legends designed for CTWR.

[5]The functional interfaces that are used to update the sound control parameters for audio indicators include parameters to indicate whether or not the map as a whole has been muted. The sound subsystem provides a separate functional interface through which the current settings of gain and pan controls for each audio indicator can be tracked by the sound subsystem. These latter controls allow a user to isolate or adjust the playback settings of each audio indicator if needed to

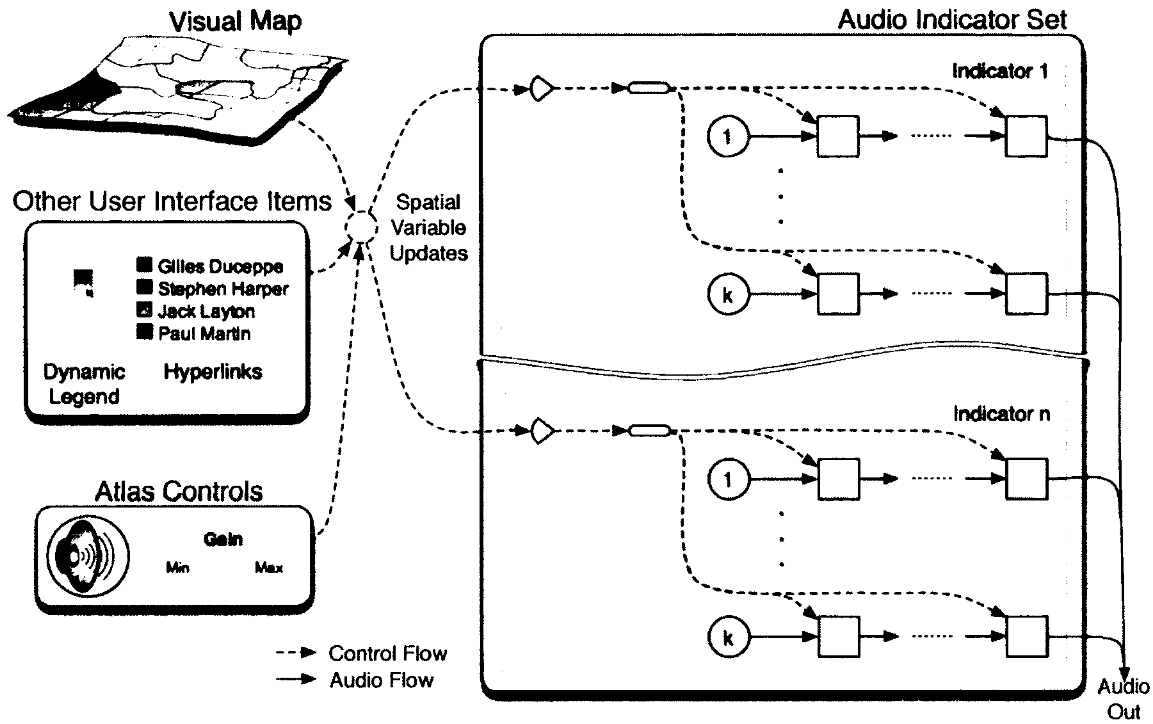


Figure 7.4: Audiovisual map sound design: Audio indicator set (figure concept and design: Glenn Brauen; the map sheet image is adapted from an image posted on http://www.openclipart.org by Ted Balmer and placed in the public domain).

A complete sound design for a map or atlas could, in addition to the linked auditory representations discussed above, contain sound components that need not be varied in accordance with a user's interactions with a map. The sound subsystem design allows for narrations or other sound components to be used without being managed through an audio indicator set. Instead, these components could be created and managed using direct access to the layered software libraries on which the audio indicator design itself is built.[6] Appropriate user interface

help with learning how the sounds behave or to concentrate on a single indicator if desired for a particular use of a map. For more information, see Appendix F.

[6]Early versions of the sound subsystem did not include the audio indicator abstraction as outlined here and all sound usage in relation to the map was managed by directly accessing this *now* lower layer of the design, called the *Sound Proxy*. Appendix F provides a more detailed description of the relationship between the audio indicator set and sound proxy designs. The sound proxy provides capabilities that enable sound files to be transferred to the client computer from the networked system hosting the map or atlas, and then allows playback of those sounds to be started,

controls to start or stop narrations or other non-varying sounds could be included as part of the map application's user interface.[7] Although these types of sound uses have been considered and designed into audiovisual map applications using the sound subsystem design described here, the focus of this section is on the use of sound components for which playback is varied based on the interactions of a user with the map.

In the sound subsystem, as currently designed, an audio source may be either sampled audio or an instrument in a Musical Instrument Digital Interface (MIDI) composition. If the audio source is sampled audio, then the audio paths shown in Figure 7.2 carry digitally encoded audio samples (i.e., arrays containing numerical representations of analog audio waveforms) and the effects processing components may use any algorithm needed to modify those samples according to the current classified value of the spatial variable.[8] For example, a gain control can lower the loudness of sampled audio by multiplying each sample value, a number representing an instantaneous waveform amplitude for the digitally encoded sound, by a value in the range $0 \leq x < 1.0$. Once processed, each sample is passed

---

stopped, or paused as required. In fact, all of the capabilities for manipulating the playback of sounds and MIDI compositions described in this section are possible through interaction directly with the lower layer libraries with the exception that the ability to drive those manipulations as a side-effect of updating monitored variables, the raison d'être of the audio indicator design, does not exist.

[7]Alternatively, it is easy to embed unlinked sounds such as narrations in web applications like the maps discussed here using commercial and widely available user interface multimedia components such as embedded QuickTime players (http://www.apple.com/support/quicktime/) and these usually include their own user interface for controlling the sound playback (i.e., play, pause, stop and volume controls). If an audiovisual map is designed to vary the playback of sounds based on the actions of a user, then running all of the sounds for a given map design, including unlinked ones, through a single sound subsystem has the advantage of allowing easier integrated control. For example, a control to mute the entire application doesn't have to do anything special for other sound players, and map sound volumes could be adjusted automatically if a narration is started.

[8]A specific set of processing algorithms has been developed to date to demonstrate the idea. The addition of new audio processing methods is possible but requires software development effort to add features to the sound subsystem.

in order to either the next effect processing component in the set or to the audio system of the computer on which the audiovisual map application is running.

If the audio source is a MIDI instrument, then the audio paths shown in Figure 7.2 may carry either MIDI performance command messages (e.g., "play middle C using medium key pressure") or digitally encoded audio samples once all of the MIDI command processing has resolved to a series of musical events generating sounds. For the purposes of this description, it is acceptable to assume that the data path for a MIDI instrument carries only MIDI messages until having passed through the last effect processing component.[9]

7.1 Example Audio Indicator Designs

This section discusses the audio indicator designs used in *Canada's Trade with World Regions: 1976–2000* (CTWR). As discussed in Appendix B, CTWR uses audio indicators to represent the following variables: *balance of trade*, and *value of exchange as a percentage of Canada's total trade*. Although the version of CTWR presented during the user evaluations discussed in Chapter 6 used a classifying version of the *value of exchange* indicator, the application includes a set of sound design selectors, some of which configure a linearly interpolated version of the indicator. For completeness of the design overview provided here, discussion of both the classified and the linearly interpolated versions of this indicator are included here to allow a discussion of their differences.

[9]When using MIDI message protocol, a synthesizer is responsible for converting note events (e.g., 'on' or 'off') to digital audio sample streams based on the current configured state of the channel being processed as described by all MIDI messages received up to the current event. In the sound subsystem design described here, a software synthesizer is used and it is possible that certain controls such as pan or output gain are applied after the conversion to digitally encoded audio samples. The model described here would apply equally to a system in which the output from the sound subsystem was passed to a hardware MIDI synthesizer and all parameter control was based on manipulation of MIDI events.

Figure 7.5 shows the effects processing configuration for the *balance of trade* variable that is updated, as a user works with the map application, with the value of the balance of trade between Canada and the region currently selected on the map. When, as a result of the cursor being moved over a new region on the map, an updated value is computed for the ratio:

$$\frac{\text{Canadian Exports to Region}}{\text{Canadian Imports from Region}} \tag{7.1}$$

the new value is passed to the classifier and it determines a value in the range $0 \ldots 2$ using the class value ranges shown. Once the correct class has been determined, it is passed to all of the effect processing components in the design (in the example shown, two classifying mute controls and two static gain controls, one of each associated with each of the MIDI instruments used in the design). The classifying mute controls use the entry in the *muteFlags* array identified by the class index to set the muting control of the associated MIDI instrument channel ('true' indicating that the instrument will be inaudible). The static gain controls actually ignore the class index because they always set the gain for the associated instrument to a preconfigured constant gain value, selected to adjust each instrument's loudness so that the two instruments are perceived as approximately equally audible (when not muted).[10]

This design produces muting control graphs for each of the two instruments used by the representation as a function of input trade balance ratio values as

[10]The adjustment of perceived loudness is dependent on parameters such as the pitch and timbre of each instrument, the note velocities (how hard each note is hit for each instrument), the sound system of the computer used, and the user's hearing. The currently configured values (guitar: 0.65, cello: 0.70 compared to a maximum gain setting of 1.0) reflect, in part, the higher register used for the melody line played on the guitar and the higher pitch of that instrument in comparison to the cello. Assuming that other factors are kept constant, a high pitch tone will tend to be perceived as louder than a lower pitch tone if both are played with the same energy.

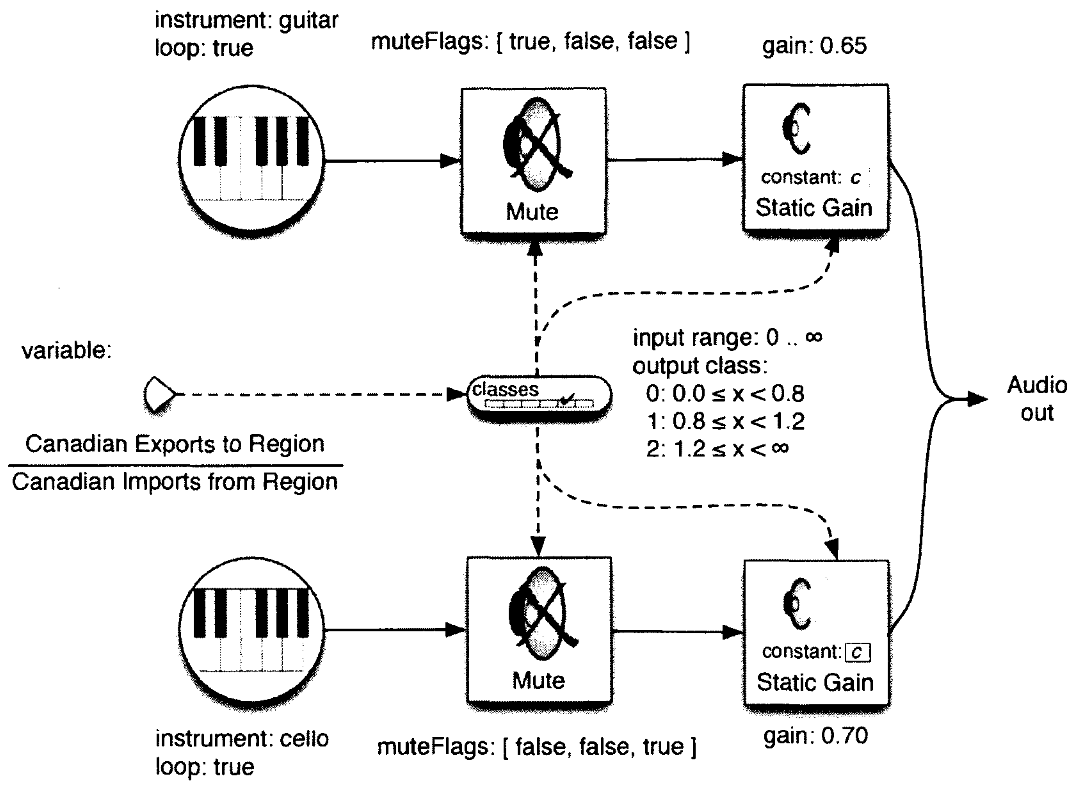


Figure 7.5: Canada's Trade with World Regions: 1976–2000 classified trade balance sound design (figure design and production: Glenn Brauen).

shown in Figure 7.6 and the muting behaviour shown for the nylon string guitar and the cello in Table B.3. When both instruments are sounding, the selected region and Canada have approximately balanced trade. When one or the other of the instruments are muted, then there is a trade imbalance and the direction of that imbalance is signified by which of the instruments is audible.

Figure 7.7 shows the effects processing configuration for both a classifying and a linearly interpolated auditory representation of the *value of exchange* variable, computed as the value of the current region's trade exchange, maintained as a percentage of Canada's total trade value.

The audio indicator design for the *value of exchange* variable uses a single MIDI instrument (a drum kit) as the sound source and adjusts the loudness of the drums

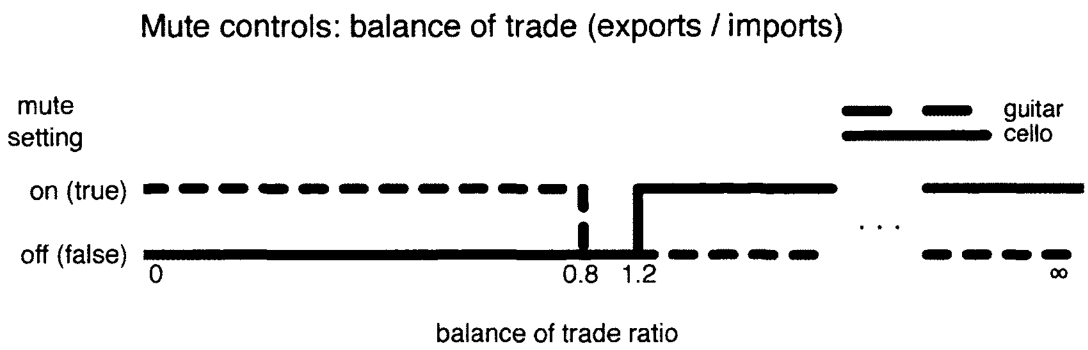


Figure 7.6: Canada's Trade with World Regions: 1976–2000 mute control graphs for instruments representing Canada's balance of trade. Note that these use negative logic: a mute setting of *true* indicates that the associated sound is *inaudible* (figure design and production: Glenn Brauen).

in proportion to the value of the region's trade with Canada. Both configurations compute the gain setting control value c based on the currently selected region's contribution to Canada's total trade computed as the following ratio (with possible values in the range of 0.0 ... 1.0)

$$\frac{\text{Value of Region's Trade with Canada}}{\text{Value of Total Canadian Trade}} \tag{7.2}$$

The classifying configuration shown in Figure 7.7a divides the range of possible values into three classes with breaks at 6.7% and 25% of Canada's total world trade and sets the gain control value c to values of 0.35, 0.675, and 1.0 for each class respectively, as shown in the *gains* array values. The linearly interpolated configuration shown in Figure 7.7b computes values of c using the equation

$$c = C_{min} + \left(\frac{v - R_{min}}{R_{max} - R_{min}} \times (C_{max} - C_{min})\right) \tag{7.3}$$

where v is the value of the ratio from equation (7.2) for the currently selected re-

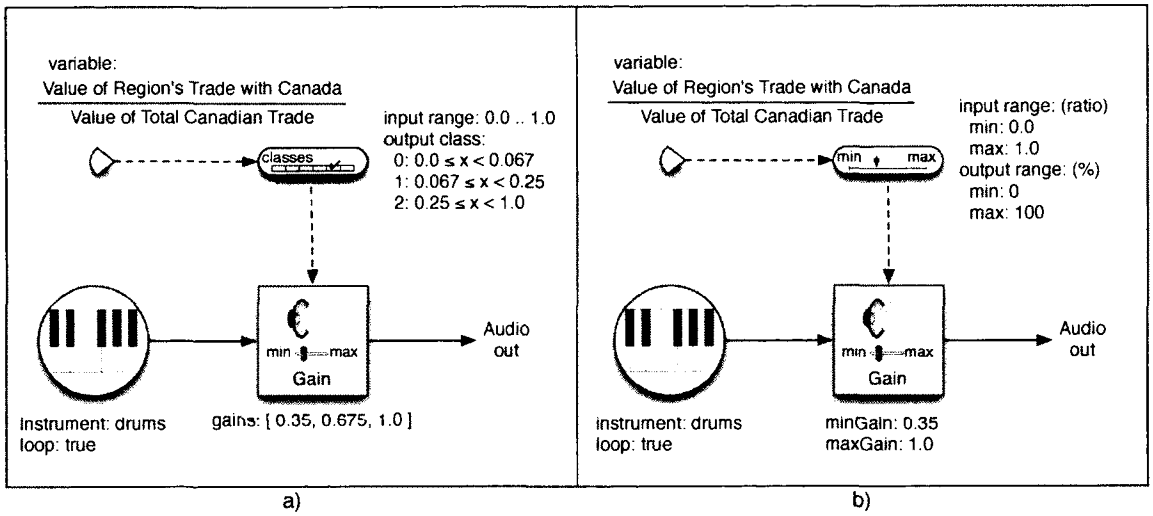


Figure 7.7: Canada's Trade with World Regions: 1976–2000 value of exchange sound design: a) classifying, b) linearly interpolated (figure design and production: Glenn Brauen).

gion, R_{min} and R_{max} are the minimum and maximum values for the variable range (0.0 and 1.0, the minimum and maximum possible ratio values respectively), and C_{min} and C_{max} are the minimum and maximum control parameter values to be used for this acoustic representation (0.35 and 1.0, the minimum and maximum gain settings in this case, respectively).

The result of either of these designs is the loudness adjusting behaviour shown for the drums in Table B.3. The loudness of the drums is controlled, for each of the classified and linearly interpolated control cases, as shown in Figure 7.8. For the classifying configuration, the loudness of the drums increases as trade value with the selected region increases but the classification differentiates between three groupings of Canada's trading partners as identified within the relatively small and highly aggregated data set used in this application. The linearly interpolated configuration adjusts the loudness in proportion to the computed *value of exchange*, although a minimum gain setting of 0.35 was configured so that the drums would

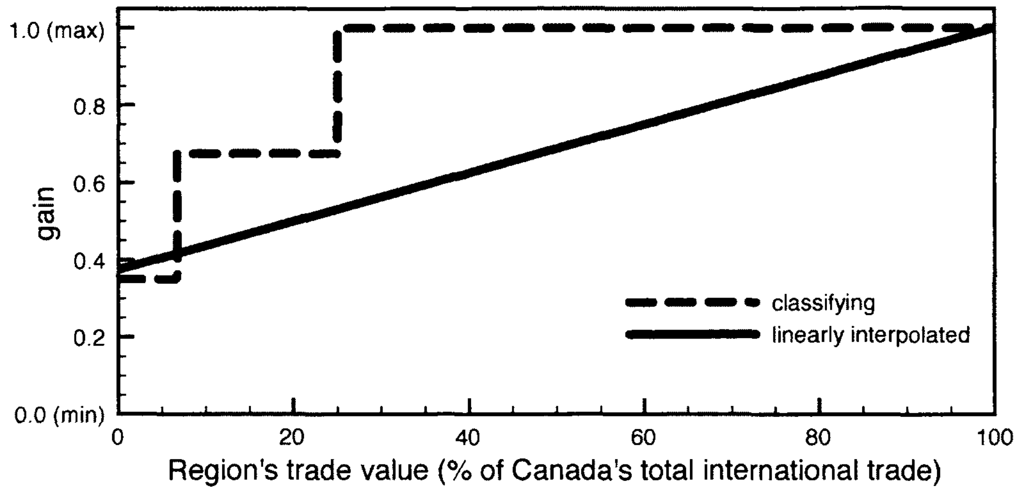


Figure 7.8: Canada's Trade with World Regions: 1976–2000 gain control graphs for audio indicators representing Canada's value of exchange (figure design and production: Glenn Brauen).

always be at least minimally audible, even for the lowest possible values.[11]

The audio indicator configuration diagrams shown in this section are, at present, only an explanatory device. As discussed in Appendix F, the software detailing the configuration of an audio indicator set is written, using Javascript (ECMA 2009), as part of the creation of an audiovisual map application. Figure 7.9 contains a listing of Javascript software that configures an audio indicator set containing classifying versions of the *balance of trade* and *value of exchange* variables, as shown in Figures 7.5 and 7.7a.

[11]The maximum values for the regional value of exchange variable are all associated with the United States which constitutes a region on its own because of the heavy reliance of Canadian trading companies on the USA as a trade source or destination. The value of trade with the USA, according to this data set when aggregated to include all trade sectors, ranges between about 69% and 77% of Canada's total value of exchange. When linearly interpolated, as described above, these maximum exchange values will result in lower gain settings than the classifying design (i.e., maximum is $0.35 + (0.77 \times 0.65) \approx 0.85$ according to equation (7.3)). To scale the upper limit on the gain control, the input range of the linearly interpolating classifier shown in Figure 7.7b could be adjusted by setting *maxGain* to the actual maximum in the data set (0.77), thereby adjusting R_{max} in equation (7.3) and producing a maximum gain setting of 1.0. Note that this implies a change to Equation (7.2) whereby the gain setting is adjusted in proportion to the greatest trade exchange value for a single region rather than Canada's total trade.

```
/*
 * The audio indicator set (aiSet) uses the identified MIDI composition for
 * all instruments.
 */
var aiSet = audioIndicatorSet("sounds/Bach_Preludio_BWV997.mid");

/*
 * 1st indicator component set - classifed 'balance' indicator.
 */
var balanceClassification = classification({ // trade balance - three levels
    limits : [
        { lower : 0.00000, upper : 0.79999 },
        { lower : 0.80000, upper : 1.19999 },
        { lower : 1.20000, upper : 9999999 }
    ]
});

var m = midiIndicator({ trackId : 1, channelId : 6 }); // nylon string guitar
m.addProcessingComponent( // add parametrized effect processors
    midiTrackClassifiedMute({
        classification: balanceClassification,
        muteFlags: [ true, false, false ] // mute only class 0
    }),
    midiChannelStaticGain({ gain: 0.65  })
);
aiSet.addIndicator('balance', m);

m = midiIndicator({ trackId : 2, channelId : 4 }); // cello
m.addProcessingComponent( // add parametrized effect processors
    midiTrackClassifiedMute({
        classification: balanceClassification,
        muteFlags: [ false, false, true ] // mute only class 2
    }),
    midiChannelStaticGain({ gain: 0.7 })
);
aiSet.addIndicator('balance', m);

/*
 * 2nd indicator component set - classified 'value of exchange' indicator.
 */
var exchangeClassified = classification({ // value of exchange - three levels
    limits : [
        { lower : 0.0,   upper : 0.0669999999 },
        { lower : 0.067, upper : 0.2499999999 },
        { lower : 0.25,  upper : 1.0 }
    ]
});

m = midiIndicator({ trackId : 3, channelId : 9 }); // drums
m.addProcessingComponent( // add parametrized effect processors
    midiChannelClassifiedGain({
        classification : exchangeClassified,
        gains          : [ 0.35, 0.675, 1.0 ]
    })
);
aiSet.addIndicator('exchange', m);
```

Figure 7.9: Canada's Trade with World Regions: 1976–2000 audio indicator set configuration.

The purpose of showing the Javascript configuration for this example sound design is to outline briefly how the software design adheres to the higher-level design described above and to provide a glimpse of the kind of design specification that currently must be created to configure the sound design for an audiovisual map. The intent here is not to provide a complete overview of how the software is used. Appendix F and the source code listings included on the CD-ROM accompanying this thesis (see Appendix G) are intended to provide a more complete description of the software design and its use.

The current requirement for an audiovisual map designer to write software to configure the sound design may prevent some people from using this without assistance. But the main objectives of this design were that it be very flexible to allow a broad range of audiovisual map prototypes to be created and tested, drawing on some of the ideas discussed in Part II of this thesis, while being organized around a reasonably understandable high-level model.

In the listing shown in Figure 7.9, an audio indicator set is created (line 5) and is passed a reference to a Stored MIDI File (SMF) containing instrumental performance descriptions to be used as sound sources for each component indicator. As discussed above, the audio indicator for the regional trade balance variable is created using two MIDI sources, a nylon string guitar performance taken from MIDI track 1, channel 6 (line 18) and a cello performance taken from MIDI track 2, channel 4 (line 28). These two instruments are controlled according to a shared classification object (created on lines 10–16) and each is created by allocating a new MIDI source and effects processing chain (lines 18 and 28). The parameter control for each instrument is configured by adding effects to mute the instrument based on the shared classification scheme (lines 20–23 and 30–33) and to assign a predetermined gain setting to each instrument (lines 24 and 34).

Once each of these source and effects processing sets are configured, they are added into the audio indicator set, each associated with the same indicator identifier 'balance' (lines 26 and 36). Subsequently, parameter adjustments for the audio indicator associated with the *balance of trade* variable will be initiated by calling the `update()` function of the audio indicator set, specifying the indicator identifier, 'balance', so that the set can update only those auditory parameter controls to which the update applies.[12]

The value of exchange indicator is created in much the same manner using a drum track defined by MIDI track 3, channel 9 of the SMF (line 49). This indicator uses its own classification scheme (lines 41–47), adds only a single classifying gain control (lines 51–54), and is added to the indicator set using the identifier 'exchange' (line 56).

The audio indicator set configuration shown in Figure 7.9 is reasonably concise but would, I think, be very difficult to understand without the high-level design of audio indicators and sets having been explained. The sound subsystem provides flexibility for defining exactly how a sound design will react to updates across the valid value range of associated variables which, for the purposes of this research, is important. The design choices available with this sound subsystem are extensive and yet the discussion so far has not nearly exhausted the possible ways in which sound could be used and manipulated as part of an audiovisual map.

As in visual mapping, the possibilities for representation include fundamental parameters of the modality employed. Where visual representations of difference can be based on visual parameters such as hue, saturation, shape, size, or pattern (or combinations thereof), each of a number of characteristics of sound, including at least loudness, timbre, pitch, tempo, the very presence of that sound (i.e., the use

[12] See Appendix F for more information.

of muting), and combinations thereof, could be the basis of representing difference using sound. Additional parameters affecting the timing, duration or ordering of how sounds are heard may also be considered (also possibly in conjunction with the characteristics above) but the linear, non-determinate manner in which a user may interact with an audiovisual map may constrain these options, thus the possibilities may be different for interactive maps and scripted map sequences.

Referring to Figure 7.9, even having decided that the basis of representation would be muting control for *balance of trade* and loudness for *value of exchange*, the designer must still decide on the details of the sound to be used (in this case, the use of music, which music to use, what instruments to use for each part); whether or not to use classification for each of the variables and, if so, how to define the class boundaries for each variable; and how to mix all of the indicators together, in this case determined by the gain settings used for each of the instruments. Quite apart from these choices, a designer must decide how to visually relate the auditory changes to the map visuals.

7.2 Sound Subsystem Assessment: Portability

The cartographic sound subsystem was implemented using a combination of Java software, running as a Java applet enabling World Wide Web browser pages to load and control playback of sampled audio or synthesized compositions, and Javascript software providing an efficient application programming interface (API) through which an audiovisual map can issue commands to the applet to manage audio (see Appendix F). These programming languages were chosen because they are supported in all commonly used web browsers and on all commonly used computer operating systems, ensuring that the software should work on most types of computers with which someone may wish to access an audiovisual map. In addition, Java Runtime Environments (JRE), the software that must be installed on a computer for it to run Java programs, have a well-developed set of software libraries providing access to the audio system built into a computer and enabling the creation of software with customizable audio processing capabilities supporting both sampled audio and MIDI synthesis. Finally, the applet as used is independent of the programming technologies used to create the visual map as long as those provide the capability for the map and web site interface components to interact with the sound subsystem through its API functions (which many programming technologies used for web development will support).

In designing these audiovisual map examples, and others as part of cybercartographic atlases produced by the GCRC, an important objective was the creation of web sites that did not require software download and installation for each map to be viewed. Additional installation steps required before a user is able to satisfactorily visit a web page reduces the number of visitors that will bother or are able, because of administrative restrictions preventing their installing new software, to use the map. For the purposes of this research, it was decided that a one-time installation of the JRE is acceptable, on the condition that additional per-map installation steps not be required (i.e., once the JRE is installed, maps are accessed in the same manner as any other web site).[13]

Sampled audio sound designs as part of audiovisual maps have been successfully distributed, accessed and used. MIDI compositions used as part of audiovisual maps require more capabilities of the computer system being used to access and reproduce the maps.[14] Despite the portability of the programming languages

[13]Often the JRE will already be installed or is also required for other applications and it is thus a minor burden.

[14]See footnote 3 on p.185.

used and although MIDI is a useful and flexible technology for composing, transmitting, and performing music, the packaging and distribution of MIDI *soundbanks*, that part of the MIDI implementation used to store the sounds used by the synthesizer,[15] varies across operating systems. Therefore, it is more difficult to ensure that a users' computer reproduces a map's sound design as intended with a MIDI-based design than it is with a digitally sampled audio design without requiring the user to do an additional installation and possibly diagnostic steps.[16]

One possible solution to this would be to define a small soundbank containing only those sounds used as part of a specific audiovisual web map and to distribute it in the same manner as any other data element used by the application (e.g., an image or digitally sampled audio file). Unfortunately, there are different soundbank file formats and no single implementation of the JRE and Java MIDI support them all. Therefore, regardless of the soundbank format chosen, there can be no guarantee that all systems accessing an audiovisual map will be able to play music using a custom soundbank. As compared to sampled audio for which both open and de facto standard formats are well known and accepted, Java MIDI remains problematic due to lack of standardization.

There are no obvious, broadly available, cross-platform, free alternatives to

[15]JRE installations for all common computer operating systems, except Microsoft Windows variants, include *General MIDI* soundbanks. To reduce JRE download sizes, those distributions that do include the soundbank contain what is termed the *minimal* soundbank, providing relatively low-fidelity sound (as compared to either a *midsize* or a *deluxe* soundbank; see http://java.sun.com/products/java-media/sound/soundbanks.html). In my work with *Canada's Trade with World Regions: 1976–2000*, I used the deluxe soundbank, mixed the instruments using that soundbank, and what I hear on my system is then unlikely to be the same as what other users would hear when using the web application. Depending on the operating system in their computer, users will by default either have the minimal soundbank or none at all, unless they have explicitly downloaded and installed a Java compatible soundbank.

[16]Many computers have built-in MIDI soundbanks as part of their audio system. Therefore, without explicitly installing the Java MIDI soundbanks with which an audiovisual map was designed the sound design may work but using different soundbanks (installed with the JRE, depending on the operating system, or built-in, depending on the computer's hardware) or may not make any sound at all.

Java and Javascript for creating a sound subsystem appropriate for use as part of audiovisual maps distributed for use over the World Wide Web that would provide all of the capabilities of the sound subsystem discussed here. Adobe ActionScript, a programming technology that can run on as wide a variety of computers and operating systems as Java and Javascript, provides capabilities for managing digitally sampled audio streams and for mixing multiple audio streams (Adobe Systems Inc. 2008, 556–583). Like the audio effects processing capabilities of the sound subsystem discussed in this thesis, Adobe ActionScript could be the basis for the development of audio processing effects (ibid.). However, Adobe Actionscript support for audio does not include support for sound synthesis (MIDI or otherwise). Finally, its use would require that computers accessing an audiovisual map built using it have a proprietary Flash browser add-on software component installed and that developers of audiovisual maps use a proprietary, commercial set of tools, predominantly designed for programming graphics applications for use as part of web sites, to create those applications, regardless of the graphics programming technologies actually used to create the visual map components.

Alternative enhanced audio processing capabilities directly supported by web browsers have been proposed (Humphrey et al. 2009; Rogers 2010) but each of these initiatives specify different mechanisms for controlling and processing sound and are each supported only in specific web browsers. Therefore, these proposals do not currently provide a basis for broadly distributable audiovisual maps.

Finally, many flexible and powerful audio processing software applications exist that could be used as part of an audiovisual map[17] but, for use with applications distributed via the World Wide Web, would require software development to inte-

---

[17]For example, MAX/MSP is one such commercial application (Cycling '74 2011) and CSound is an open source project (Boulanger 2000).

grate them with web browsers (i.e., either inter-process communication interfaces or application programming interfaces) to enable them to respond to requests to manage audio in conjunction with a user's interactions with a map. These solutions would require that each computer accessing an audiovisual map application have additional software installed, again creating an installation hurdle potentially reducing use of audiovisual web maps.

—

The software for the cartographic sound subsystem described in this chapter was designed to support modular definition of audio indicators, each composed of a sound source and an ordered set of effect processing components, all controlled by a variable monitor and a classifier, for the following reasons: 1) clarity, 2) flexibility within a single map sound design, 3) reuse, and 4) potential for visual programming. First, using this approach it is possible to produce reasonably clear descriptions of each audio indicator and of their aggregate behaviour. Second, this approach recognizes that multiple audio indicators used within a single map should not be constrained such that they all use the same or similar designs (e.g., all varying gain or all using classified variables) and allows for detailed parameter control once a metaphor for acoustic representation (e.g., *louder means more*) has been selected. Third, this approach recognizes the importance of being able to easily reuse and recombine audio effect processing components, each of which has a simple clearly defined function, in part, because the inter-related perception of sound parameters (e.g., pitch and loudness) means that multiple parameters will often need to be controlled together within the design of an audio indicator. Fourth, this modular approach could be the basis of a visual configuration system that would potentially alleviate the need for direct Javascript programming, thereby allowing people without Javascript programming skills to define audio indicators.[18]

As was stressed with respect to film and game sound designs, the range of sound design possibilities for an audiovisual map are almost limitless and to date, to the knowledge of this author, cartographic research into this range has not been conducted. As Taylor (2005c) argued, ideas concerning cybercartographic practices develop through combined theory and practice. In that spirit, this sound subsystem was refined and extended based on an earlier design[19] in parallel with the development of a new audiovisual map example. Chapter 8 describes that application and discusses some of the design ideas examined through these paired processes.

[18]The development of an authoring tool to generate the audio indicator configurations remains for future research. To be really useful, this type of an authoring tool would need to be incorporated into a larger configuration tool for audiovisual maps and designing such an end-user tool that alleviates the need for programming remains a challenge.

[19]See footnote 1 on p.181.

8

An Example Audiovisual Map Application: Canada-USA Commodity Trade 1976–2000

This chapter describes an example audiovisual map application created to demonstrate design ideas concerning interactive and animated audiovisual cartography as part of this thesis. In particular, design issues concerning the sound design and the staging of multiple audio indicators as a group and in relation to the visual map updates caused by user actions are discussed.

Canada-USA Commodity Trade 1976–2000 (CUCT)[1] was created using a database concerning trade between Canada and the United States of America. The data was acquired from the International Trade Division of Statistics Canada by researchers developing maps and data sets for the *Cybercartographic Atlas of Canada's Trade with the World* (CACTW).[2] This chapter describes CUCT only as required to discuss its sound design. This description may be clarified, if required, by examining the application available on the World Wide Web,[1] or videos showing the application being used (available on the CD-ROM accompanying this thesis; see Appendix G).

[1] Available at http://atlas.gcrc.carleton.ca/tdb/canusa/. The map shown in Figure 8.1 is from the page design found by following the *Interactive Map* or *Sequenced Maps* links. Videos showing the maps being used are included on the CD-ROM enclosed with this thesis (see Appendix G).

[2] High level descriptions of the data are provided by Eddy (2005a, b).

In addition, Appendix H provides a more complete overview of the application, showing page layouts, providing functional descriptions of user selectable options, and providing complete definitions of the acoustically represented variables and designs for the audio indicators.

Two of the example applications discussed earlier in this thesis because of their inclusion in the user evaluations described in Chapter 6, *Ottawa Area Federal Election Sound Map: June 28, 2004* and *Canada's Trade with World Regions: 1976–2000*, were designed to be consistent with the objective of testing users working with auditory representations of data for which no correlated visual display was provided.[3] This objective limited to some extent the focus on the visual design of these examples.[4] Therefore, although the user evaluations showed both positive and negative responses to the map examples and the design ideas evaluated, I thought scope existed to more thoroughly examine *integrated auditory and visual design* and to that end CUCT was developed to fulfill the following objectives: to create an example with a more balanced approach to both the visual and auditory elements of the integrated design; to create an example based on a more extensive data set that could be the basis of future user evaluations, including repeated tests with individuals over time to assess longer-term usability and usefulness,[5] as opposed to only looking at the initial encounters users have with a map application as was done in the tests discussed in Chapter 6; to be used as a test case for the con-

---

[3]This is not true of the audiovisual examples from the *Cybercartographic Atlas of Antarctica*. These examples were designed to fulfill a variety of objectives within the scope of the atlas project.

[4]As discussed in Appendix B, *Canada's Trade with World Regions: 1976–2000* was modified for use in the evaluation sessions to further limit the visual design of the application, removing the bar chart displays and options for selecting trade sector subsets or reporting periods.

[5]*Canada's Trade with World Regions: 1976–2000* was created using highly aggregated data, showing continental trade groupings and five year periods, matching the data aggregation used in an earlier, silent, prototype developed for the *Cybercartographic Atlas of Canada's Trade with the World* (CACTW). When I met with members of Statistics Canada's Geography Division, Brad Ruth expressed interest in the audiovisual design but argued that as aggregated it provided a very limited view of the data and would be difficult to evaluate for effectiveness.

current redesign of the cartographic sound subsystem, described in Chapter 7; and to explore ideas concerning *sequenced* (or *animated*) maps in addition to interactive maps, following the discussion of Chapter 5.[6]

Canada-USA Commodity Trade 1976–2000 allows a user to study international commodity trade between Canada and the United States, providing audiovisual representations of annually aggregated trade and allowing selection of an overview of all trade sectors or, as shown in Table 8.1, a further refinement to examine sub-categories of each of three top level commodity type classes: *Materials and Energy*, *Food and Agriculture*, and *Manufactured Goods*.

Regardless of the configured selection criteria, CUCT represents trade using a visual trade flow diagram and a set of audio indicators. The visual trade flow diagram, as shown in Figure 8.1, consists of a set of pie charts each scaled in proportion to the value of trade between the associated region and the region currently selected on the map (Québec in Figure 8.1), based on configured user options. The pie charts are linked to the currently selected region by flow lines indicating which of imports or exports are being displayed. Legends show the currently displayed trade sector classification, according to a user selected option from the top- or second-level sector categories shown in Table 8.1, and pie chart sizes for a selected set of trade values spanning the range of values in the database.

Audio indicators are used to represent: 1) the *value of exchange* for the currently selected region, showing its contribution to total Canada-US trade (as a percentage of the total); 2) the *balance of trade* for the currently selected region; and 3) the rel-

[6]In CUCT, what are now referred to as *sequenced* maps were originally called *animated* maps, consistent with previous cartographic research (Campbell and Egbert 1990; Peterson 1995; Thrower 1959). As will be discussed, to better accommodate the sound design I decided to animate all map transitions using a simple form of frame filling, regardless of the cause of the transition. Therefore, the map displays in CUCT use a simple form of animation for both the interactive map page and for the sequenced maps page in which preconfigured criteria determine the ordering with which trade data flows are displayed for the user.

Table 8.1: Canada-USA Commodity Trade 1976–2000 sector hierarchy. Designed for *Cybercartographic Atlas of Canada's Trade with the World* (Eddy 2005a, b).

All Sectors		
Minerals and Metals	Staple Foods	Industrial Machinery
Oil, Gas and Coal	Fruits, Grains and Vegetables	Bulk Transport Goods
Wood, Pulp and Paper	Livestock and Meat	Automotive Goods
Chemicals and Materials	Fish and Seafood	Special Equipment
Electrical Power		Household Goods
		Clothing and Fabrics

ative *value of $1* (Canadian, current) adjusted for inflation, only in the sequenced maps and if the *adjust for inflation* option is selected.[7] The *value of exchange* and *balance of trade* variables are spatial and are computed for the current trade data selections, including sector, year, flow direction, and region.[8] The inflation adjusted *value of $1* is a function only of the current year and the option to make it audible is especially useful with map sequences showing a region's trade over time.[9]

The audio indicators for CUCT were designed to complement the visual display by voicing relative measures not readily visible in any single view of the data. The value of the current region's exchange as a percentage of Canada-US trade does not appear as part of any single map display, each of which shows only the absolute values of the flows for the current region (including the textual displays showing total Canada-US trade for the current region). The current region's balance of trade can be examined visually only by changing the selection of trade flow

[7]All trade flows values are adjusted according to Statistics Canada's Consumer Price Index (CPI) using 1976 as the base year. The absolute value legend shown on the map is redrawn based on the observed range of inter-region trade flow values adjusted for inflation.

[8]The current value of these variables are selected by cursor location in the interactive map along with the trade data selection options for year, trade sector, and flow direction (exports or imports). In the sequenced maps page, the current value of these variables is defined by preconfigured sequence type, trade data selection options, and the current state of a running sequence.

[9]One CPI adjustment is applied to trade values for all regions based solely on the year selected. Therefore, there is no spatial component to the CPI adjustments and this will not affect the comparison of concurrent trade flows for different regions.

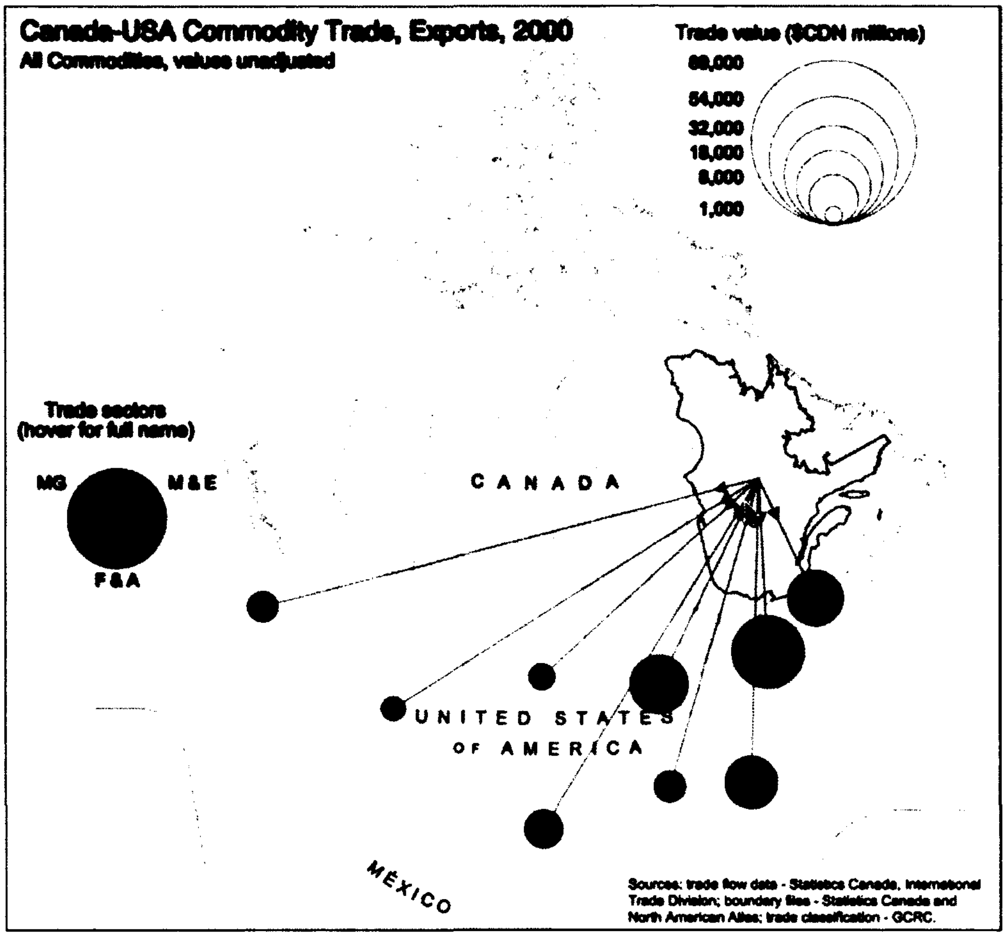


Figure 8.1: Canada-USA Commodity Trade 1976–2000 visual trade flow display for top level trade sector exports from Québec during 2000. Design and Production: Glenn Brauen. Trade classification: *Cybercartographic Atlas of Canada's Trade with the World* (Eddy 2005a, b). Trade data: Statistics Canada, International Trade Division. Boundary files: Statistics Canada and North American Atlas.

direction in the application's selection panel and then redisplaying the data for the same region (or in the sequenced maps page using the import/export sequence type). Using an audio indicator, by comparison, allows the relative values of the import and export flows to be represented concurrently as a related measurement for the selected region. The inflation adjusted value of $1 is only voiced as an accompaniment, when selected, for map sequences and is not readily apparent in the application otherwise. This audio indicator is intended as a reminder of the chang-

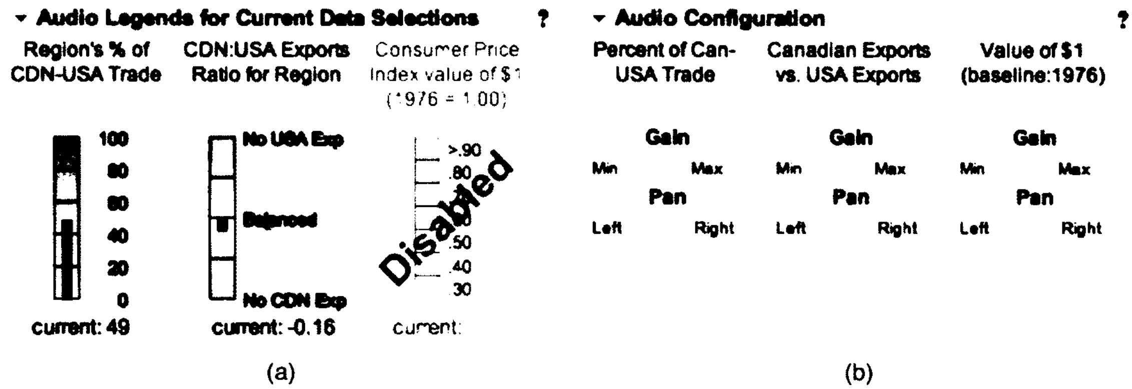


Figure 8.2: Canada-USA Commodity Trade 1976–2000 visual design components supporting auditory indicators in the sequenced maps page: a) dynamic audiovisual legends, and b) audio configuration controls (figure concept and design: Glenn Brauen).

ing basis of comparison over time, when inflation adjustment is selected, and as a correlative variable to be compared against temporal trends in trade value.

Figure 8.2 shows dynamic audiovisual legends and audio configuration controls designed as part of CUCT. Intended as learning aids to assist users in understanding the sound design, the meters can be used as direct audio controllers by positioning the cursor over the meter causing the application to adjust the sound for that indicator according to the variable value associated with the cursor location.[10] They are also updated, along with the sound design, to display current variable values as shown in Figure 8.2a as a result of cursor location on the map. If the *adjust for inflation* option is not selected, the associated dynamic audiovisual legend is disabled, as is shown in Figure 8.2a. Figure 8.2b shows a set of slider controls for gain and pan that can be used to adjust each audio indicator for listening preferences, peculiarities of computer sound systems, or when a user wants for any reason to focus on a subset of the available audio indicators.

The visual and interactive design of CUCT was influenced by the earlier de-

[10]Section B.2 provides a more detailed description of the use of the classifying dynamic audiovisual legends used in *Canada's Trade with World Regions: 1976–2000* (Figure B.3).

sign of *Canada's Trade with World Regions: 1976–2000* (CTWR), the use of dynamic audiovisual legends in that application, and the results of the user evaluations of CTWR (Chapter 6). From the outset, the design of CUCT supported the acoustic representation of quantitative data using dynamic audiovisual legends. Because of the increased density of data selection and configuration controls in the CUCT application (trade sector, year, and flow direction selectors; textual display of trade data in the interactive map page; sequencing configuration controls in the sequenced maps page), it was unclear if the dynamic audiovisual legends could be permanently visible within the application. Based on the user evaluations of CTWR, especially the result indicating that, at least during initial use, participants preferred to have the dynamic audiovisual legends visible, and based on informal user testing of early CUCT prototypes,[11] the legends have been designed to be visible by default although they share screen space with the audio configuration controls shown in Figure 8.2. To accommodate this placement of the legends in the CUCT application interface, they have been made smaller. In addition, because of user responses to CTWR indicating that training dramatically improved users' abilities to consider the sound design as quantitatively meaningful and to understand the sound design, I have added explanatory videos accessible through help hyperlinks in the map application interface explaining, among other things, the fact that these are audiovisual maps and the use of the dynamic audiovisual legends.[12] Finally, the CUCT application opens on an explanatory overview page that discusses many aspects of the design, again intended to assist in initial explanation and training for users visiting the web site for the first time, within a networked

[11]Sébastien Caquard suggested that I try to make the legends visible all the time after testing an early prototype.

[12]The help hyperlinks are shown as question mark ('?') icons in Figure 8.2. These videos are also included on the media and source code CD-ROM packaged with this thesis.

map distribution model. Only informal usability testing for CUCT, such as that mentioned above on early prototypes, has been conducted at the time of writing.

The remainder of this chapter highlights issues concerning the sound design of CUCT.

8.1 Sound Design Issues

Having experimented with the use of synthesized music to represent trade data in *Canada's Trade with World Regions: 1976–2000* (CTWR), the sound design for CUCT was created using recorded audio. This approach was used for two main reasons. First, based on user evaluation feedback showing the tendency among participants to not interpret the music in CTWR as signifying trade data, a sound with a stronger metaphorical relation to trade flows was desired. Second, problems with the packaging and distribution of the Java implementation of Musical Instrument Digital Interface (MIDI) capabilities used by the the cartographic sound subsystem, as discussed in Section 7.2, reduced the predictability and hence the effectiveness of MIDI sound design transcoding.

The predominant sound in CUCT is provided by the *value of exchange* audio indicator. This sound plays continuously while the map is actively in use, either because a user has positioned the cursor over a region on the interactive map or because a user has configured and started a map sequence running using the sequenced maps page. The other audio indicators for *balance of trade* and, when in use, the inflation adjusted *value of $1*, sound briefly to mark value changes of each associated variable but are otherwise silent.

The *value of exchange* audio indicator loops a composite recording that emulates the sound of a train speeding up or slowing down, controlled by the ratio of the value of the current region's Canada-US trade compared to the total, creating a design based on the metaphor 'faster means more.' A gain control is used, reducing the gain setting slightly as the playback speed of the recording increases, to counteract the perception of increased loudness caused by the higher frequencies produced when the recording is sped up. The loop contains a mix of the recording of a cargo train at a level crossing along with a second recording containing higher frequencies, resulting in a composite sound that becomes less of a wash of indistinct white noise at low playback rates caused by low values being assigned to the variable.[13]

The *balance of trade* audio indicator combines the playback of two audio clips, balancing the gain settings of the two to indicate which of Canadian exports or American exports are predominant according to user configured trade options (sector, year, and region). Originally conceived as a trade balance indicator for the map of Canada's trade with the world (Chapter 7 and Appendix B.2) and adapted from that for use in the map of Canada-US trade, various types of sound and control parameters were used to compare a region's exports with its imports (Figure 8.3a): music with different registers and tempos associated with each directional trade flow; brief vocal narrations expressing a classified measure of the relative difference between the directional flows;[14] and abstract sounds with combinations of pitch, pan, and gain settings adjusted according to the relative value of each directional flow. Although some of these approaches created designs in which changes in trade balance caused distinguishable sound changes, these designs were very

[13] Recording credits and discussion of variants tried are provided in Appendix H.

[14] In theory, synthesized vocal narrations could be used without classification. I chose not to use voice synthesis because I find it is still not as pleasant to listen to as a real voice, because I would have had to create more software to use this approach, and most importantly because with the size of the data set available for *Canada-USA Commodity Trade 1976–2000* it was not clear that classification could be applied effectively. To use recorded vocals, a predetermined set of phrases must be prepared which necessitates the use of classification.

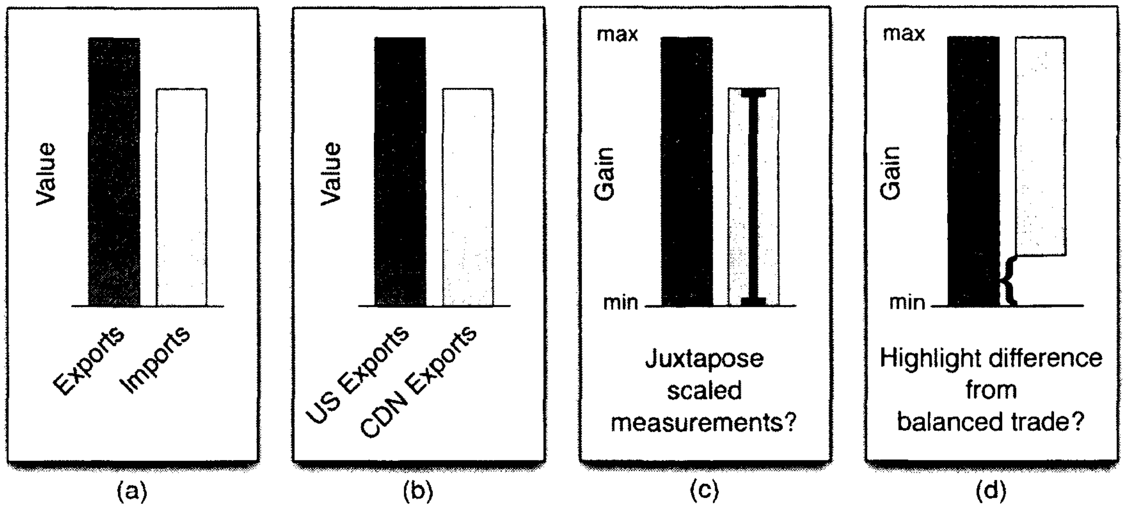


Figure 8.3: Alternative approaches to comprehending and representing a region's balance of trade in the context of *Canada-USA Commodity Trade 1976–2000*: a) comparing a region's imports versus exports; b) comparing exports from Canada to exports from the United States for the region; c) juxtaposing scaled indications of two variables; and d) representing the difference between two variables (figure design and production: Glenn Brauen).

abstract with little that metaphorically related the sounds to trade flows. The difficulty of linking a sound metaphorically to trade flows was increased by the changing polarity of trade flows depending on the identity of the currently selected region because exports may flow from Canada to the United States or vice versa depending on the country in which the current region is located.

The current design of the *balance of trade* audio indicator uses two audio recordings in a symmetric manner, a steam burst associated with Canadian exports to the United States and a computer start-up chime associated with US exports to Canada.[13] This represents the same computed values as shown in Figure 8.3a but reconceptualizes them as paired export flows (Figure 8.3b), one from Canada and the other from the United States, meaningfully reframing the data in the case of a bilateral data set such as is used in *Canada-USA Commodity Trade 1976–2000*. Discarding import/export value pairings in favour of Canada/US export pairings, with the concomitant removal of trade flow polarity changes discussed above, allows for more specific sounds to be selected to represent each trade flow. The computer start-up chime is suggestive of the predominance of manufactured goods being exported from the United States to Canada while the steam burst could be suggestive of the greater proportion of primary goods and energy in exports from Canada to the USA. Although uses of sound such as this could caricature national trade relations, reducing a complex set of transactions to a single acoustic association in each direction, this approach may support a more memorable directional association for each sound, once learned.

Having decided on acoustic associations to be used when comparing the value of the directional trade flows, additional choices remain concerning how to express such a comparison. A visual bar graph of the values to be compared works by juxtaposing scaled representations of each, as shown in Figure 8.3c, and an acoustic analogy for this would be to play each sound concurrently, adjusting an acoustic parameter such as gain according to the value of each associated flow. This approach was tried, resulting in both the steam burst and the start-up chime playing whenever the trade variables were updated, setting one sound to maximum gain while the other was (gain) scaled to represent a relatively lower trade value. In the current design, an alternative scheme was used whereby the trade flows are compared and the difference between the two, computed as a ratio, is expressed by playing the sound associated with the larger flow, scaling that sound's gain in proportion to the difference between the value of the two flows as shown in Figure 8.3d.[15] This design emphasizes variances from balanced trade. Small absolute valued trade flow ratios are voiced using a quiet rendition of the sound associated with the greater flow while a greater absolute valued ratio is represented as

[15]See Equation (H.2) in Appendix H for details concerning this ratio computation.

a louder rendition of the sound for the predominant flow (Figure H.5). Therefore, this design attempts to attract user attention more strongly to greater trade imbalances, as reflected in the computed ratios.

The *value of $1* audio indicator divides the possible range of inflation-adjusted values into classes representing $0.10 ranges, beginning with class 0 defined as $\$0.30 < x \leq \$0.40$ and continuing up to class 6 defined as $x > \$0.90$. Each class is associated with a brief narration that voices that value range and is controlled such that it is played only if the underlying variable is updated to a value that falls within the class range.[16]

The audio indicators for the *balance of trade* and *value of $1* variables use hysteresis to prevent overly repetitious playing of each associated sound. *Balance of trade*, represented without classification, uses a threshold to prevent very small value changes compared to the last value heard from being expressed. *Value of $1*, classified for representation, only voices changes that result in the selection of a different class than that of the last value heard. These uses of hysteresis are intended to reduce tedious repetition of sounds that indicate little change from previous values and to attempt to only draw user attention when a significant change, rather arbitrarily determined, is detected. However, this design could potentially cause a user to simply stop considering the variable associated with an indicator for which a sequence of variable updates deemed insignificant have been muted.

The *balance of trade* and *value of $1* audio indicator designs also delay the start of audio playback for each associated sound to create a sequence of signifying sounds, along with the continuously playing audio representing the *value of trade* variable, expressing the updates for all three of the variables as a group. This discussion next considers the staging of the audio indicators, taken as a group and

[16]See Appendix H for a detailed discussion of the design.

in relation to the visual map updates.

8.1.1 Staging Map Updates: Audiovisual Update Timing

The initial design of the visual map created static trade flow images (see Figure 8.1). When a user moved the cursor from one region to the next, the display would be redrawn to shown the trade data for the newly selected region. The confluence of the trade flow lines would move in a single update from (approximately) the geographic centre of the previously selected region to the (approximate) centre of the newly selected region. As will be discussed below, one of the difficulties with initial design attempts to accommodate multiple audio indicators in this map centred on the difficulty of creating an understandable and acceptably pleasant representation of the trade data that was also brief enough that the auditory events caused by an update would not seem noticeably delayed compared to the visual map updates. This has been achieved through efforts to improve the sound design complemented by changes in the behaviour of the visual map that grant some slack allowing the time necessary for the auditory updates to be played.

Audiovisual maps *perform* spatial data. This is true of *animated* maps (*sequenced* maps, as I have been referring to them as will be explained in this section) and to some extent of interactive maps but not to the extent that it is true of the sounds in audiovisual maps.

As has been briefly mentioned in the preceding overview of the audio indicator designs used in *Canada-USA Commodity Trade 1976–2000*, timing considerations are apparent throughout the designs. What are the durations of audio recordings being used? Is an audio recording looped continuously? When is an audio recording played relative to its triggering event? If acoustic parameters are being adjusted, how long a fade is used to adjust the control? In the case of some acoustic parameter controllers, such as playback rate controls, how does it affect the duration and timing of the audio being controlled? If multiple audio indicators are being used together, how are the various sounds staged for combined expression?

This list of questions focuses predominantly on the time characteristics of sounds used in a design, rather generically, without much consideration given to the types of sounds used and the implications of those for the audiovisual map as a performance. Chapter 4 discussed possible functions of sound at length, including uses of specific types of sounds. Therefore, I will restrict this discussion to a consideration of timing issues without further consideration of sound types.

In *Canada-USA Commodity Trade 1976–2000*, the audiovisual map can be controlled in two different ways: through direct user interaction or through preconfigured map sequences. The choice of control mode (through the selection of the interactive map or sequenced maps page) further affects the possibilities for altering the performed timing of the audiovisual map sound design. Figure 8.4 shows the timing of key acoustic events for the sound design of *Canada-USA Commodity Trade 1976–2000*, relative to an update event causing value changes in the spatial variable set underlying the sound design, when used in either the interactive map or the sequenced maps page, with each possible sequence speed shown for the latter. For each of these cases, Figure 8.4 shows the control parameter transition timing for the *value of exchange* audio indicator as a colour change in the solid line representing the continuous play of the looping train audio recording along with marks indicating when each of the *balance of trade* and *value of $1* audio indicators may be audible. Because of update hysteresis for each of the intermittently heard indicators, as discussed above, they may not always be audible as shown in the diagram if an update has not altered the previous variable value by a large enough

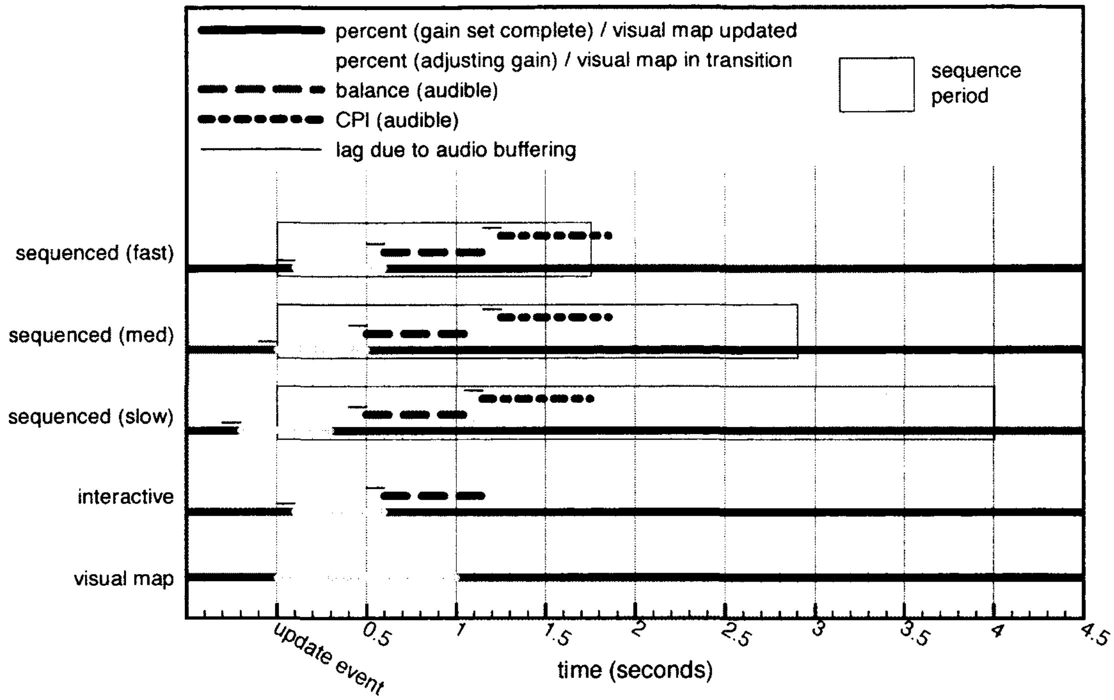


Figure 8.4: Canada-USA Commodity Trade 1976–2000 audiovisual update timings for visual map, interactive audio indicators, and sequenced audio indicators (figure design and production: Glenn Brauen).

amount to warrant repetition. Also shown are update lags of approximately 100 ms in the audio processing subsystem caused by audio buffering intended to even out the transmission of audio data to the computer's audio system to account for small delays that can be caused by network transmission delays (in the case of streaming audio) or lack of responsiveness in the map user's computer itself.[17]

While using the interactive map, the *value of exchange* and *balance of trade* audio indicators are updated as a result of a user moving the cursor across a mapped

[17]When last examined, the output buffering caused approximately 100 ms of lag on average, predominantly caused by that amount of audio information already having been queued toward the audio output of the computer before an update event such as a cursor movement is processed, the current value of audio control parameters are changed, and altered audio begins to be added to the same audio output buffer. Assuming that an audio stream is being transmitted over a very responsive network and is being processed and played using a very responsive computer, the output buffers can be shortened to reduce this lag but the risk exists that the transmitted stream will not keep up with the required output audio data rate, at least intermittently, causing audio to stall.

region boundary, initiating the recomputation of the spatial variables and adjustment of the audio control parameters. There is a delay before any change begins to be heard caused by the buffering lag (approximately 100 ms), then the gain and playback rate parameters for the *value of exchange* audio are adjusted incrementally over a period of 500 ms,[18] causing a noticeable transition in the audio playback speed. The *balance of trade* audio indicator design delays playback of the associated audio clips by 500 ms after the initiating event is detected, resulting in a delay of approximately 600 ms before the sound for this indicator begins to be heard because of this delay combined with the buffering lag.[19] The combined delays cause the *balance of trade* sounds to play only after the *value of exchange* audio indicator has completed its transition. A similar delay before the initiation of playback is used with the *value of $1* audio indicator in the sequenced maps to delay the associated vocal narrations until after the *value of exchange* audio indicator has completed its control parameter transitions and the *balance of trade* audio indicator has fully played its recordings.[20]

By staging the sounds over time as shown in Figure 8.4, the combined designs of the audio indicators are an attempt to allow each indicator to be heard without interference from concurrent (significant) acoustic events. A potentially negative effect of this staging is the increased duration of the period over which acoustic events resulting from an update event occur. In the design of *Canada-USA Commodity Trade 1976–2000*, and particularly in early attempts to stage only the two audio indicators for the interactive map, there seemed to be a tension between the sudden single visual update abruptly moving the confluence of trade flows from one region to another and resizing all of the proportional trade value sym-

[18]This delay is configured as part of the design shown in Figure H.2.
[19]This delay is configured as part of the design shown in Figure H.4.
[20]This delay is configured as part of the design shown in Figure H.6.

bols before the acoustic updates had even begun. The combination appeared to emphasize the delay in the audio indicators causing me to attempt various strategies to reduce the duration of the acoustic update period with the result at least sometimes of creating a sound design that still lagged noticeably behind the visual update but was also composed of sounds that felt rushed. Reducing sound durations too much would also risk making acoustic parameters of those sounds, such as loudness, more difficult for a user to judge thereby reducing the signifying effectiveness of the audio indicators.

In retrospect it is not clear that the length of time being allowed for the acoustic updates to occur was on its own the problem but rather the contrast between the timing of the visual and acoustic updates. After many attempts and upon deciding that I was not going to be able to satisfactorily reduce the apparent acoustic lag as compared to an instantaneous visual update, I decided to reverse the problem by animating the transition of the visual map to its newly updated visual state over time. With experimentation, I settled on an animated transition of 1 second, during which the flow lines reorient to a newly selected map region and the new proportional symbols are faded into view about halfway through the visual transition. This design, as seen and heard in the current version of the application and as shown in Figure 8.4 as a colour change in the line representing the continuous visibility of the map, still completes the visual transition before the acoustic update period but both now occur in a comparable time. The result of this redesign is an audiovisual map design that seems more relaxing to work with.

The introduction of animation to better align the transition timing of the visual and auditory components of the map may be an example of the type of integrated design Théberge (2005) was suggesting when he argued that sound should be considered more than an add-on to an existing visual map. As an extension of this, the algorithm used to animate the visual map updates (line motions and fades, proportional symbols fades) was conceived and written as a straightforward analogy of the earlier algorithm used to change audio playback parameters (e.g., gain, playback speed) over a specified period of time within the sound subsystem. Therefore, in this research, not only has design work been completed to ensure that the visual map accommodates the sound design but lessons learned in producing sound designs have been applied to change the behaviour of the visual map.

Because this visual animation procedure was applied to both the interactive and sequenced maps pages, the term *sequenced maps* was adopted for what I had before then referred to as the *animated maps*. The interactive map has become an hybrid, incorporating reactive animation for transitions triggered by user actions and needing to respond to new update events even if the last update cycle has yet to complete. The visual transitions in conjunction with the acoustic updates seem to make the map easier to follow as the animated motion of the converging trade flow lines draw a user's attention toward the next selected region more gradually rather than requiring a shift of focus after the lines have moved in an instantaneous update.

In the interactive map, the next update event, determined by cursor movements, is unknown until it has happened. Therefore, in Figure 8.4, the buffering lag is marked as beginning when the update event is detected and processing of the new acoustic parameter updates begins. This is not the case in the sequenced maps for which a sequence type and associated options have been preconfigured by a user before they press the *Start Sequencing* button. For these, not only is the timing of the next update known but the next set of values for the spatial variables are also known, allowing adjustments to be made to the timing of the acoustic update periods as shown in Figure 8.4 and Table 8.2.

Table 8.2: Canada-USA Commodity Trade 1976–2000 audio indicator update event timing adjustments.

	Timing offsets, by audio indicator (milliseconds)		
map (speed)	**Value of exchange**	**Trade balance**	**Value of $1 (CPI)**
sequenced (fast)	0	0	0
sequenced (med)	-100	-100	0
sequenced (slow)	-300	-100	-100
interactive	0†	0†	n/a‡

† No timing offset possible because timing of next update event is unknown.
‡ Audio indicator for the inflation adjusted value of $1 is unused with the interactive map.

For example, when a map sequence configured on the slowest speed setting is running, the staging of audio indicators is relaxed to provide more time around each and the transition of the *value of exchange* audio indicator is shifted to commence before and continue into the period during which visual map updates are proceeding. This is accomplished by starting the updates for the *value of exchange* and *balance of trade* audio indicators 300 ms earlier than is done for the interactive map. Taking the output buffering lag into account, this causes the *value of exchange* audio indicator's transition to start about 200 ms before the visual map update begins. This advance is noticeable if you listen and look for it and, as is often done in film scene cuts and with the use of sound cues for which the associated actor or object in a game is not yet visible, the sound transition can be used to anticipate what is about to happen with the visual map once the sound design has been learned. By comparison, using the medium sequencing rate, the timing of the audio indicators is adjusted only to compensate for the output buffer lag, resulting in the visual and acoustic updates seeming to be synchronized. The fastest speed setting of the sequenced maps does not adjust the timings in comparison to the interactive map.

Figure 8.4 also shows the sequence update periods, the delay between trade data observation display updates, for each of the configurable sequencing rate options. For the slowest rate (4 seconds), the map's acoustic and visual update period[21] is about 2 seconds or half of the sequence update period. For the medium rate (2.9 seconds), the acoustic and visual update period is approximately 1.85 seconds. For the fastest rate (1.75 seconds), if the *value of $1* audio indicator is heard, then the acoustic and visual update period is actually longer that the sequence update period and runs into the beginning of the next update cycle.

Different staging times for the acoustic and visual map updates in comparison to the sequence update period in use result in different perceptual experiences when working with the map. The slowest sequence rate seems quite relaxed with the map being static approximately half of the time such a sequence is running. The medium sequence rate has a noticeably quicker pace but the audiovisual experience still seems well coordinated with noticeable pauses during each sequence update period. By contrast, the fastest sequence rate seems somewhat relentless and at times it becomes difficult to maintain context while using a sequence. I almost always run sequences using one of the two slower sequence rates but kept the fastest setting for demonstration purposes.

In this discussion, I highlight these timing issues to emphasize the audiovisual map as a performance of spatial data. Although many questions concerning the usability of static maps remain open, the performance of animated and audiovisual maps increases the potential variability of the map use experience and will require study. I am unaware of literature discussing update timing for interactive or animated maps and introduce the topic concerning audiovisual maps here.

The hybrid construction of the interactive map, incorporating simple frame filling animation to coordinate the visual and acoustic updates further highlights the

[21]From the beginning to the end of the time during which all perceptual acoustic and visual updates occur, assuming that *adjust for inflation* is selected and the update plays the audio for the *value of $1* audio indicator, the latest staged of the audio indicators.

map's performance of a selection of trade data. Although no rigorous user evaluations have yet been done with this map nor with the configurable sequence rates supported, the difference of the audiovisual map's *feel* when played using the different rates seems obvious. How any of these settings would affect user enjoyment, comprehensibility, engagement and learning would require further study.

9

Discussion and Conclusions

Maps and mapping processes, under a variety of labels (e.g., the terms 'map' and 'geographic visualization'), are changing. A wide variety of types of maps, and many of them capable of incorporating multiple media types, are used to address a diversity of user groups, topics, requirements, and use contexts in which spatial information and media conveying that information can support decision making. But as argued throughout this thesis, uses of sound as part of maps and spatial media have received little attention as a subject of cartographic research. Research agendas for cartography still show strong visual biases and a recent statement concerning research agendas for the International Cartographic Association stated that "[m]aps and geographic information have special power through their ability to connect and integrate data sets by the inherent geographical location, and present the information contents in a user-friendly and understandable *visual* and *tactual* way" (Virrantaus et al. 2009, 63, emphasis added). The possibility of using audio interfaces was mentioned only with respect to usability research concerning map design for visually impaired users (ibid., 67).

In this thesis, I have attempted to integrate design ideas for sound with dynamic visual design ideas for maps, have developed a theoretical framework within which audiovisual cartography could be studied, have created a model for implementing particular audiovisual map applications, and demonstrated a software implementation of a sound subsystem meeting the requirements of that model in the context of web-based interactive audiovisual cartography. To guide this research, I have reviewed literature concerning sound design for film and games, and literature from human-computer interaction (HCI) research concerning uses of sound for education and as an alternative modality for visually impaired people. As part of the process, applying theory in parallel with developing it, I have created and collaborated in the creation of several audiovisual map examples both as a basis for discussion of sound design ideas for audiovisual cartography and as test materials to be used for user evaluations. I conducted user evaluations, in collaboration with HCI researchers, that generated data concerning the usability and usefulness of specific maps but also provided a basis for limited discussion of some of the audiovisual design ideas proposed by this research.

As was discussed in Chapter 1, most map designs currently found on the World Wide Web that incorporate sound do so by organizing audio spatially according to attributes associated with each sound (e.g., location of recording) rather than having been constructed to incorporate a specific sound design as a thematic signifying element of the map itself. Extending the approach of using the map as an organizational method for disparate sound recordings, online web maps will sometimes incorporate user contributed sounds into the map (e.g. Stein 2008), further separating the visual and acoustic design processes used in the creation of the map with a coordinated visual design used to merely *collect* the acoustic elements. Similarly, as was discussed for the classification of "sound maps" shown in Tables A.1 and A.2 (pages 269 and 270, respectively), the user interaction models often separate sound playback from the visual map interface by making the user click through to audio players that are visually separated from the map. The map remains a silent entity. Although all of these approaches are valid, the lack of consideration of sound as a design element for interactive mapping, both in the cartographic literature and in the June 2010 survey of online map examples discussed in Section 1.3, can be interpreted as indicating a bias among people making maps, and in the tools supporting map design activities, when compared to the prevalence of sound in a range of other media and the purposes for which sound is deployed in those media.

One contribution of this thesis then is to highlight the discrepancy between the possibilities for which cartographic sound design could be applied compared to the ways in which it is commonly used. Within the framework for the creation of meaning using sound in audiovisual cartography discussed in Chapter 3, Chapter 4 discussed motivations for the use of sound in audiovisual cartography. Chapter 5 proposed a set of functions for which sound could be appropriately deployed as an alternative modality complementing the visual design of dynamic maps. The range of possibilities for sound design as part of interactive audiovisual cartography is potentially great and has yet to be extensively researched. Even the examples discussed in this thesis, those designed by me and those from the *Cybercartographic Atlas of Antarctica* on which I collaborated, although demonstrating a range of reasons for and methods of including sound in the designs and in the types of sounds used, have only scratched the surface of the possibilities for audiovisual design for dynamic maps.[1]

[1]By contrast, the development and deployment of sound in cinema and games were and are ongoing projects and have experienced considerable change and experimentation in both the purposes for which sound is used and in the technological means employed to convey sound within cinema. Despite ubiquitous presence in contemporary cinema, ways and means of employing sound in cinema were and are the focus of much economic, technological, and artistic activity over long periods beginning in the nineteenth century, continuing through the twentieth and into the twenty-first century (for perspectives on the development of sound in American cinema, see

This thesis has also made a contribution in considering the technological problems associated with integrating the use of sound into an audiovisual map to be used as a World Wide Web application. Chapter 7 outlined the high-level design of a sound subsystem intended to be used for the creation of interactive audiovisual map applications. The sound subsystem was redesigned following the premise that creative audiovisual map design requires a very flexible model that can support a range of sound design approaches and that can be quickly reconfigured to try new ideas during map creation. Currently, the sound subsystem is a stand-alone programming toolkit which enhances its flexibility, allowing it to be integrated with different visual mapping technologies. Although this enhances its technical flexibility, it makes use of the toolkit possible only within project teams that can manage the complexities of the required programming. Although it could be beneficial to some potential users to provide an easy-to-use configuration interface for the sound subsystem, such an approach inevitably increases maintenance costs for the overall system. This thesis has, after consideration of these issues, simply outlined a possible visual approach to the creation of such an end user configuration interface while leaving the research to create such an interface, and to integrate it into an encompassing mapping tool, as a future research challenge.

The cartographic sound subsystem discussed here supports the use of two types of input acoustic specifications: digitally sampled sound and synthesized musical (MIDI) compositions. Each of these two forms have been successfully used in audiovisual map designs discussed in this thesis. From a research perspective, sound designs based on each of these two types of sound inputs offer

Gomery 1976, 2005; Stewart 1980). Over a much shorter and more recent period, the history of musical composition and sound design for computer games can only be described as one of constant change affecting all stages in the production, distribution, and performance of games and their sound designs (Collins 2008b; Sanger 2003).

alternative possibilities and warrant further study. Practical considerations concerning the difficulty of ensuring accurate reproduction of sound designs based on Java MIDI by the variety of computer types that may be used by people to access an audiovisual map over the World Wide Web make this choice less attractive, as discussed in Chapter 7. In addition, user models for digitally sampled audio, based as they are on consumer products such as tape recorders, are much more familiar to many potential map designers than are the more specialized and complex MIDI technologies. The comparative simplicity of digitally sampled audio would normally not matter to users directly.[2]

In parallel with the sound subsystem redesign described in Chapter 7, specific sound design problems were examined in the construction of the audiovisual map application *Canada-USA Commodity Trade 1976–2000* as discussed in Chapter 8. The application incorporates sound effects and narration to convey summary data complementing the synchronized visual display of trade information. But, as discussed, the application's visual design also accommodates the sounds used, providing dynamic audiovisual legends and configuration controls to assist users in learning and understanding the sound design. In addition, the application's visual map updates are animated to provide time for the sound design transitions used to signify changes in the summary variables.

More noticeably than is the case with silent interactive and animated maps, the design of an audiovisual map highlights presentation timing issues, possibly because of the greater sensitivity of human hearing than sight to event timing. These

[2]Unless audiovisual map designers, for example, deciding to take advantage of libraries of music residing on the same computer with which a user is accessing an audiovisual map application and selected according to that user's individual tastes, offered options allowing users to select their own music as part of an hybrid audiovisual map / jukebox design. In such a case, not necessarily that improbable when considering some of the types of hybrid, multimedia maps that have appeared in recent years, relatively few users would have MIDI compositions at hand and the greater complexity of MIDI file structures would make their use more error prone.

timing issues may influence the reception and understanding of the map by users and additional research concerning this is required. The timing of visual map updates relative to acoustic updates, possible timing delays between separate audio indicators, and all of these in relation to the actions of a user may impact interpretation of the information. As discussed in Chapter 6, information represented aurally such as using speech is intrinsically temporal and managing this information in response to non-linear, indeterminate user actions can also create synchronizations of auditory and visual information that are difficult to predict and may cause unexpected interpretations by users.

For many subjects for which an audiovisual map could be developed, there will be no 'correct' sound to use and designing a sound to signify a specific concept may be a challenge. For the exploration map from the *Cybercartographic Atlas of Antarctica*, what does a satellite sound like? For *Canada-USA Commodity Trade 1976–2000*, what does a trade balance sound like?[3] Depending on audience and the objectives for a map, examining sound designs from other media may provide useful ideas. Similar to the many visual design choices made in the process of constructing a map, understanding the necessary decisions made while designing sounds for an audiovisual map highlights the status of that object as a consciously constructed representation.

Map use is a complex process that, as Carter (2005) argued, encompasses multiple dimensions including characteristics of map users and their reasons for consulting a specific map, the uses for which a specific map is intended, communities within which a map may be used, the circumstances within which a map is pre-

[3]Even when recording an event in the world is possible, sound designers for narrative films often invent sounds, for artistic and economic reasons, to be paired with visuals (Sonnenschein 2001, 57–58). Authenticity may be an objective, especially with documentary, but may be counterbalanced with the desire for effective communication.

sented and used, and the specifics of the design of a map being used. Of these topics, the introduction of auditory signification into maps impacts most directly on the design of the map itself but it also constrains the circumstances of map use because of the requirement to maintain components of the map as digital media throughout the authoring, distribution, and use of the map. Other dimensions of map use apply to audiovisual maps in much the same way as they apply to maps generally. When engaging with media, possibilities for engaging sensory modalities are influenced by material capabilities of a medium and the technological, economic, and social understandings of the medium by all those in contact with it, producers and users alike (Gitelman 2006; Jenkins 2006; van Leeuwen 2005). Users' expectations for sound as part of an audiovisual cartography will be shaped, at least in part, by their experience with sound in other media, if only because of the relative paucity of sound use in cartography.

Use of an audiovisual map on its own is multimodal but, if used in conjunction with other media with which a user may be engaged, introduces the map into a potentially complex simultaneous mix of media. A map user is as free as any other media user to mix and modify media and turning off a map's sound design to listen instead to something else is always an option. Thus the sound design for a map, if it is to be listened to, must be understood *by the user* to be adding value to the combined audiovisual experience. Maps, as opposed to cinema and computer games, are not (*yet*) widely considered to be auditory artefacts which means that the presence of sound may be more likely to be noticed and critically appraised than when used in a context in which the use of sound has been naturalized (e.g., cinema and games). Ideally, understanding the context within which a particular map or atlas is used should always be part of the analysis of the requirements for the project and should receive attention during iterative design and testing cycles.

This thesis has also contributed to discussions concerning methods for validating the appropriateness of design choices concerning, in this case, audiovisual cartography through user evaluations although the method used here would also be appropriate for interactive or animated cartography. Although the results from usability testing can't be generalized beyond discussion of those systems tested, the results discussed in Chapter 6 demonstrated that the interpretation of visual and auditory map components in the examples tested was inter-related for some users. Furthermore, if the intent of the sound designs were adequately understood or explained, users were often able to work with and explain information presented using auditory or combined audiovisual representation. The results also demonstrated that these users mostly accepted sound as part of the application designs, although self-selection bias must be considered because the advertisements for participants clearly indicated that map applications using sound were to be tested. However, the test results also showed mixed results concerning the effectiveness of the particular designs tested: participants spent time trying to understand the purpose of the sounds; trying to isolate a particular sound when designs intentionally mixed multiple sounds, especially recordings of speech; and participants were sometimes unclear if the information they required to answer questions was available in the audio or visuals designed into the applications (or was available in both or neither). Therefore, although it appears that the use of sound in maps changes their use and interpretation, it also seems that there is much yet to be learned about the design options available and their effects on map use and understanding.

Taking together the difficulties of generalizing results from usability testing of audiovisual map examples and the many possibilities for the use of sound remaining to be tried within specific mapping projects, consideration must be given to determining efficient and useful methods for understanding which approaches to sound design seem appropriate for the characteristics of a particular mapping project (e.g., age, gender, and experience level of users along with thematic foci of the project and the materials available from which maps can be created). As was discussed in Chapter 6, the in-depth qualitative analysis of user evaluation sessions conducted in this research was useful in allowing participant verbalizations and the visual and acoustic elements of the audiovisual map examples being examined to be related and to attempt to do so using a small enough time scale that immediate reactions to events during the evaluation session could be examined with respect to what the participant was doing and how the application was responding. Sometimes apparently contradictory reactions to a map example were thus possible to explain as reactions to separate elements of the design occurring in sequence.

Although useful, this detailed qualitative analysis was also very time-consuming and would be difficult to use except sparingly as part of a larger cartographic design process because of the reporting delay after testing. Fortunately, data can be generated during evaluation sessions in such a way that alternative methods of analysis can be applied, making it possible to postpone the decision of which methods to use until after data collection or to analyze subsets of the data differently. As with silent maps and other computer applications, project-specific design activities (e.g., user needs analysis including interviews and questionnaires) and testing activities addressing quantitative and qualitative aspects of usability and usefulness as part of iterative processes focused on project objectives and representative user groups are probably most economical in gauging the effectiveness of all aspects of an audiovisual map application against focused application criteria. Perhaps directed use of more detailed qualitative analysis methods, such as that used here, when particular questions arise within the analysis of data produced during the evaluation sessions would be one way in which to focus the method and limit the time spent.

General guidelines for sound design in audiovisual maps may emerge over time at least for certain application areas[4] and among certain groups of users. However, design differences are likely to persist due to varying project objectives, implementation technologies, and preferences among users and designers. It would be surprising if this were not the case, considering the wide variety of possible design approaches for silent interactive and animated maps. Many questions concerning designs for audiovisual map applications remain open and it would be premature to prescribe design guidelines for the use of sound or sound types in audiovisual maps. The design of effective audiovisual maps requires that many issues concerning the integration of sound and visuals, the design of user interfaces, and the pacing of multimodal information presentation be evaluated to ensure that users are able to achieve their objectives. User expectations for sound design will continue to develop in response to existing and new ways in which they experience sound being used, because of ongoing transformations in the spaces in which they encounter music and other composed sounds,[5] and because of their increasingly mobile use of both auditory media (Bull 2000, 2007) and maps within those spaces (e.g., navigation systems and location-based services).

If, as Sui and Goodchild (2001) argued, one of the primary roles emerging for

[4]For example, indigenous place name pronunciation maps are emerging as one application of audiovisual maps (Gwich'in Social and Cultural Institute 2003; Mouafo and Müller 2002; Nunavut Youth Consulting et al. 2008), although the interaction models vary from one to the next.

[5]Schafer (1977, 103–109) described changes in concert music during and after the industrial revolution and argued that the compositional inspirations for music, the structure of concert halls, the organization and size of orchestras, and the types and sound production characteristics of the instruments they used all changed as a result of the increasingly urban environments in which composers and musicians lived. More recently, recording technologies, digital technologies, and widespread adoption of portable audio devices continue to modify approaches to composition and how people listen to music (Cox and Warner 2004).

GIS is the communication of spatial information for the general public, then the prevalence of sound in many media intended to reach wide audiences would seem to argue for ongoing research concerning the communicative potentials of maps and GIS, in at least some of which a focus on uses of sound would be appropriate. However, the role of human-computer interaction in selecting data for sonification as part of the model of audiovisual map use discussed by this thesis makes this model highly appropriate for use in private, exploratory map uses as well as for public presentations (MacEachren 1994).

As is the case with many forms of computer-based cartography, audiovisual cartography currently requires a diverse set of skills. This is so partly because of the range of subject matter knowledges and technical skills required to design, create, and deploy contemporary maps and mapping applications (Cartwright et al. 2007; Peterson 2008; Taylor 2008; Taylor and Pyne 2010) but also because well-developed tools and processes required to design and integrate auditory media into audiovisual maps don't exist, or are not mature and stable, or change quickly. The range of capabilities required of a team planning to undertake an audiovisual cartographic project to be distributed and used over digital networks would likely include, at least: subject matter knowledge, geospatial data management, cartography and interactive visual design knowledge, audio recording and processing, music composition and processing, software development, and network and database management and deployment. Transdisciplinary skills are required and the introduction of music composition, sound design and production skills are an incremental addition to this.[6]

Whereas it is almost unthinkable that a contemporary film or computer game

[6]Within research projects, Rice and Rice (2001) argued that, rather than cartographers attempting to become composers and musicians, there are many students of composition or working musicians that could be brought into such projects.

would be developed without the inclusion of a sound design, silent maps are still the default and indeed the question of adding sound is probably almost never consciously considered. The economics of map-making are very different from those of at least some film and game development projects, and in many contexts the time and cost required to consider audiovisual design for a mapping project may still argue against it. At the same time, since the introduction of the Sony Walkman, if not since earlier audio technologies such as the transistor radio, mobile audio has been changing the relationship between people and the media they use (Bull 2000, 2007; McLuhan 1964; Tacchi 2002). Changes in peoples' relationships to and understanding of auditory media, and changes in the networked media channels through which people access and use all media, including maps, may provide opportunities for new map forms and map hybrids to be developed and tested, thereby providing some justification for taking on the additional complexity of audiovisual design for maps, at least in some projects. The release of the cartographic sound subsystem developed during this research as a freely available open source toolkit, appropriate for use in audiovisual map applications distributed and used over the the World Wide Web, is intended to alleviate one obstacle to the use of sound (Appendices F and G).

Bibliography

12 Gates to the City. 2009. 12 gates to the city: Acoustic map [web-map application]. Available at http://12gatestothecity.com/extras/acoustic-map. Last accessed June 2010.

Adobe Systems Inc. 2008. Programming Adobe ActionScript 3.0 for Adobe Flash. Updated February 11, 2009.

Agarwal, Anshu and Andrew Meyer. 2009. Beyond usability: Evaluating emotional response as an integral part of the user experience. In *Proceedings of the 27th International Conference on Human Factors in Computing Systems*, 2919–2930. New York, NY: ACM.

Altman, Rick. 1992a. Section introduction: Four and a half film fallacies. In Altman (1992d), 35–45.

———. 1992b. Section introduction: Sound/history. In Altman (1992d), 113–125.

———. 1992c. Sound space. In Altman (1992d), chap. 2, 46–64.

Altman, Rick, ed. 1992d. *Sound Theory / Sound Practice*. New York: Routledge.

Altman, Rick, McGraw Jones, and Sonia Tatroe. 2000. Inventing the cinema soundtrack: Hollywood's multiplane sound system. In *Music and Cinema*, eds. James Buhler, Caryl Flinn, and David Neumeyer, 289–297. Hanover, NH: University Press of New England.

Amundsen, Roald. 1912. *The South Pole: an account of the Norwegian Antarctic expedition in the "Fram," 1910-1912*, vol. II. London: J. Murray. Translated by Arthur G. Chater.

Aristotle. 1931a. On sense and the sensible. In *The Works of Aristotle*, vol. 3. Oxford: Clarendon Press. Translated by John Isaac Beare. Available as an eBook in the ublic domain.

———. 1931b. On the soul. In *The Works of Aristotle*, vol. 3. Oxford: Clarendon Press. Translated by John Alexander Smith. Available as an eBook in the public domain.

Asche, Hartmut and Christian M. Herrmann. 1994. Designing interactive maps for planning and education. In MacEachren and Taylor (1994), chap. 12, 215–242.

Baddeley, Alan D. 1986. *Working Memory*. New York: Oxford University Press.

Baecker, Ronald M. 1995. *Readings in Human-Computer Interaction: Toward the Year 2000*. San Francisco, Calif: Kaufmann.

Bal, Mieke. 1985. *Narratology: Introduction to the Theory of Narrative*. Toronto: University of Toronto Press. Translated by Christine van Boheemen.

Ballas, James A. 1994. Deivery of information through sound. In Kramer (1994a), chap. 2, 79–94.

Barthes, Roland. 1967. *Elements of Semiology*. New York: Hill and Wang. Translated by Annette Lavers and Colin Smith. Originally published in 1964.

———. 1972. *Mythologies*. London: J. Cape. Selected and translated by Annette Lavers. Originally published in 1957.

Baulch, Samantha, Ronald MacDonald, Peter L. Pulsifer, and D. R. Fraser Taylor. 2005. Cybercartography for education: The case of the cybercartographic atlas of Antarctica. In Taylor (2005b), chap. 21, 491–515.

BBC. 2009. BBC save our sounds [web-map application]. Available at www.bbc.co.uk/worldservice/specialreports/saveoursounds/index.shtml. Last accessed June 2010.

Becker, Brett. 2005. Puget soundscape [web-map application]. Available at http://homepage.mac.com/brettbecker/soundscape/pugetsoundscape.html. Last accessed June 2010.

Belazs, Bela. 1945. Theory of the film: Sound. In Weis and Belton (1985), 116–125.

Berg, Charles Merrell. 1976. *An Investigation of the Motives for and Realization of Music to Accompany the American Silent Film, 1896-1927*. New York: Arno Press.

Bergson, Henri. 1988. *Matter and Memory*. New York: Zone Books. Translated by N. M. Paul and W. S. Palmer.

Blattner, Meera M., Denise A. Sumikawa, and Robert M. Greenberg. 1989. Earcons and icons: Their stucture and common design principles. *Human-Computer Interaction* 4:11–44.

Block, Ned. 1981. *Imagery*. Cambridge, MA: MIT Press.

Bolter, Jay David. 2007. Remediation and the language of new media. *Northern Lights* 5:25–37.

Bolter, Jay David and Richard Grusin. 1999. *Remediation: Understanding New Media*. Cambridge, MA: MIT Press.

Booth, Heather and Jacob Brancasi. 2008. Open sound new orleans [web-map application]. Available at www.opensoundneworleans.com/core. Last accessed June 2010.

Boulanger, Richard, ed. 2000. *The CSound Book: Perspectives in Software Synthesis, Sound Design, Signal Processing, and Programming*. Cambridge, MA; London: MIT Press.

Brauen, Glenn. 2006. Designing interactive sound maps using Scalable Vector Graphics. *Cartographica* 41(1):59–71.

———. 2011. Interactive audiovisual mapping: BTEX emissions from NPRI reporting facilities in Montreal. In *Mapping Environmental Issues in the City: Arts and Cartography Cross Perspectives*, eds. William Cartwright, Sébastien Caquard, and Laurene Vaughan, chap. 6, 74–108. Berlin; Heidelberg: Springer.

Brauen, Glenn, Stephanie Pyne, Amos Hayes, Jean-Pierre Fiset, and D. R. Fraser Taylor. 2011. Encouraging transdisciplinary participation using an open source cybercartographic toolkit: The atlas of the Lake Huron Treaty relationship process. *Geomatica* 65(1):27–45. Canadian National Report to the International Cartographic Association.

Brave, Scott and Clifford Nass. 2003. Emotion in human-computer interaction. In *The Human-Computer Interaction Handbook*, eds. Julie A. Jacko and Andrew Sears, 81–96. Hillsdale, NJ, USA: L. Erlbaum Associates Inc.

Brewster, Stephane A., Peter C. Wright, and Alistair D. N. Edwards. 1994. A detailed investigation into the effectiveness of earcons. In Kramer (1994a), chap. 19, 471–498.

Bull, Michael. 2000. *Sounding out the City: Personal Stereos and the Management of Everyday Life*. Oxford; New York: Berg.

———. 2007. *Sound Moves: iPod Culture and Urban Experience*. New York: Routledge.

Burkholder, J. Peter. 2006. A simple model for associative musical meaning. In *Approaches to Meaning in Music*, eds. Byron Almén and Edward Pearsall, chap. 5, 76–106. Bloomington; Indianapolis: Indiana University Press.

Butler, Judith. 1992. Contingent foundations: Feminism and the question of "postmodernism". In *Feminists Theorize the Political*, eds. Judith Butler and Joan W. Scott, chap. 1, 3–21. New York: Routledge.

Butler, Toby and Steve Whiting. 2005. Memoryscape audio walks [web-map application]. Available at www.memoryscape.org.uk. Last accessed June 2010.

Buziek, Gerd. 2000. Legend designs for noninteractive cartographic animations. *Computers and Geosciences* 26(1):21–28.

Callard, Andrea, Fred Krughoff, and Johnny Farrow. 2006. audioBus [web-map application]. Available at www.audiobus.org. Last accessed June 2010.

Campbell, Craig S. and Stephen L. Egbert. 1990. Animated cartography: Thirty years of scratching the surface. *Cartographica* 27(2):24–46.

Cantizzani, Juan, Mariam Caballero, Miguel Angel Lastra, Sergio Millán, Pablo Sanz, Justin Bennet, and Marcel Cobussen. c2006. Andalucía soundscape [web-map application]. Available at www.andaluciasoundscape.net/mapa. Last accessed June 2010.

Caquard, Sébastien. 2009. Foreshadowing contemporary digital cartography: A historical review of cinematic maps in films. *The Cartographic Journal* 46(1):46–55.

Caquard, Sébastien, Glenn Brauen, Benjamin Wright, and Paul Jasen. 2008. Designing sound in cybercartography: From structured cinematic narratives to unpredictable sound/image interactions. *International Journal of Geographical Information Science* 22(11):1219–1245.

Carpenter, Edmund. 1973. *Eskimo Realities*. New York: Holt, Rinehart and Winston.

Carroll, Noël. 1988. *Philosophical Problems of Classical Film Theory*. Princeton, NJ: Princeton University Press.

Carter, James R. 2005. The many dimensions of map use. Paper read at XXII International Cartographic Conference (ICC2005), at A Coruña, Spain.

Cartwright, William and Michael P. Peterson. 2007. Multimedia cartography. In Cartwright et al. (2007), chap. 1, 1–10.

Cartwright, William, Michael P. Peterson, and Georg Gartner, eds. 2007. *Multimedia Cartography*. Berlin: Springer-Verlag, 2nd edn.

Castells, Manuel. 2001. *The Internet Galaxy: Reflections on the Internet, Business, and Society*. Oxford; New York: Oxford University Press.

Casti, Emanuela. 2000. *Reality as Representation: The semiotics of cartography and the generation of meaning*. Bergamo, Italy: Bergamo University Press.

Cavalcanti, Alberto. 1939. Sound in films. In Weis and Belton (1985), 98–111.

Centeno, Hector and Don Sinclair. 2008. Mississauga sound map [web-map application]. Available at www.yorku.ca/caseaces/soundmap. Last accessed June 2010.

Chen, Shing-Ling S. 1998. Electronic narcissism: College students' experiences of Walkman listening. *Qualitative Sociology* 21(3):255–276.

Chicago Public Radio. 2007. Chicago sound map [web-map application]. Available at www.chicagopublicradio.org/Soundmarks_Browse.aspx. Last accessed June 2010.

Chion, Michel. 1992. Wasted words. In Altman (1992d), chap. 5, 104–110.

———. 1994. *Audio-Vision: Sound on Screen*. New York: Columbia University Press.

City Lore and Local Projects. 2002. City of memory [web-map application]. Available at www.cityofmemory.org/map/index.php. Last accessed June 2010.

Classen, Constance. 1993. *Worlds of Sense: Exploring the senses in history and across cultures*. London: Routledge.

———. 2005. The witch's senses: Sensory idealogies and transgressive feminities from the Renaissance to modernity. In Howes (2005), chap. 4, 70–84.

CLNS and OMCA. c2003. Listening to nature: A sound walk across California [web-map application]. California Library of Natural Sounds and Oakland Museum of California. Available at http://museumca.org/naturalsounds/home.html. Last accessed June 2010.

Coburn, Craig A. and A. William Smith. 2005. Musical landscapes using satellite data. In *SPARK Festival of Electronic Music and Art. 3rd Annual Conference*. University of Minnesota, Minneapolis, MN.

Cohen, Michael and Elizabeth M. Wenzel. 1995. The design of multidimensional sound interfaces. In *Virtual Environments and Advanced Interface Design*, eds. Woodrow Barfield and Thomas A. Furness III, chap. 8, 291–346. Oxford; New York: Oxford University Press.

Collins, Karen, ed. 2008a. *From Pac-Man to Pop Music: Interactive Audio in Games and New Media*. Hampshire, UK; Burlington, VT: Ashgate.

Collins, Karen. 2008b. *Game Sound: An Introduction to the History, Theory, and Practice of Video Game Music and Sound Design.* Cambridge, MA; London: The MIT Press.

Corbin, Juliet M. and Anselm L. Strauss. 2008. *Basics of Qualitative Research.* Thousand Oaks, California: Sage, 3rd edn.

Cordobés, Mar and Richard Picanyol. 2008. Badiafonia [web-map application]. Available at www.badiafonia.net. Last accessed June 2010.

Cosgrove, Denis E. 1998. *Social Formation and Symbolic Landscape.* Madison, Wisconsin: University of Wisconsin Press.

Cox, Christoph and Daniel Warner. 2004. *Audio Culture: Readings in Modern Music.* New York: Continuum.

Cox, Trevor. 2010. Sound tourism: A travel guide to sonic wonders [web-map application]. Available at www.sonicwonders.org. Last accessed June 2010.

Crampton, Jeremy W. 2002. Interactivity types in geographic visualization. *Cartography and Geographic Information Systems* 29(2):85–98.

Crang, Mike. 2003. The hair in the gate: Visuality and geographical knowledge. *Antipode* 35(2):238–243.

Créatec +. 2005. Reading and buying books for pleasure: 2005 national survey final report. Government of Canada Publication CH44-61/2005E-PDF, Canadian Heritage. Available at http://dsp-psd.pwgsc.gc.ca/Collection/CH44-61-2005E.pdf. Last accessed February 9, 2010.

Cycling '74. 2011. MAX/MSP [software]. Available at http://cycling74.com/. Accessed August 2011.

Dant, Alastair, Tom Davis, Victor Gama, and David Gunn. c2006. Folk songs for the five points [web-map application]. Available at www.tenement.org/folksongs/client. Last accessed June 2010.

Dant, Alastair, Tom Davis, and David Gunn. 2007a. Cinco cidades [web-map application]. Available at www.cincocidades.com/en/soundmap. Last accessed June 2010.

———. 2007b. Manchester: Peripheral [web-map application]. Available at www.manchesterperipheral.com/phase1/soundmap. Last accessed June 2010.

de Lange, Catherine. 2010. Audio zoom picks out lone voice in the crowd. *New Scientist* 2781:24.

Dein, Alan and Francesca Panetta. 2010. Sound map: The Caldeonian Road [web-map application]. Available at www.guardian.co.uk/society/interactive/2010/apr/26/caledonian-road-sound-map. Last accessed June 2010.

DiBiase, David, Alan M. MacEachren, John B. Krygier, and Catherine Reeves. 1992. Animation and the role of map design in scientific visualization. *Cartography and Geographic Information Systems* 19(4):201–214.

Doane, Mary Ann. 1985. The voice in the cinema: The articulation of body and space. In Weis and Belton (1985), 162–176.

Dorling, Daniel. 1992. Stretching space and splicing time: From cartographic animation to interactive visualization. *Cartography and Geographic Information Systems* 19(4):215–227, 267–270.

Downs, Roger M. and David Stea. 1977. *Maps in Minds: Reflections on Cognitive Mapping*. New York: Harper & Row, Publishers.

Dransch, Doris. 1995. *Temporale und Nontemporale Computer-Animation in der Kartographie*. Ph.D. thesis, Freie Universität Berlin, Berlin.

Driver, Felix. 2003. On geography as a visual discipline. *Antipode* 35(2):227–231.

Drori, Gili S. 2005. *Global E-litism: Digital Technology, Social Inequality, and Transnationality*. New York: Worth Publishers.

———. 2010. Globalization and technology divides: Bifurcation of policy between the "digital divide" and the "innovation divide". *Sociological Inquiry* 80(1):63–91.

Dyson, Frances. 1995. In/quest of presence: Virtuality, aurality, and television's gulf war. In *Critical Issues in Electronic Media*, ed. Simon Penny, chap. 2, 27–46. Albany, NY: SUNY Press.

Eason, Peter and Jonathan Lewis. 2007. Soundscape of China [web-map application]. Available at www.pbs.org/kqed/chinainside/soundmap/index.html. Last accessed June 2010.

ECMA. 2009. Standard ECMA-262: ECMAScript language specification. European Computer Manufacturers Association. Available in PDF format at: http://www.ecma-international.org/publications/standards/Ecma-262.htm. Accessed October 2010.

Eddy, Brian. 2005a. Ammendment to geocarto model document. Geomatics and Cartographic Research Centre discussion paper.

———. 2005b. A geographic-cartographic model (gcm) for trade atlas content development: Considerations for implemenation in the cybercartographic atlas of canada's trade with the world. Geomatics and Cartographic Research Centre discussion paper.

Eddy, Brian G. and D. R. Fraser Taylor. 2005. Applying a cybercartographic human interface (CHI) model to create a cybercartographic atlas of Canada's trade with the world. In Taylor (2005b), chap. 22, 517–540.

Elections Canada. 2004. Federal election results: 38th general election, june 28, 2004. Government of Canada, available at www.elections.ca. Accessed March 2005.

Erevelles, Sunil. 1998. The role of affect in marketing. *Journal of Business Research* 42(3):199–215.

Ericsson, K. Anders and Herbert A. Simon. 1993. *Protocol Analysis: Verbal Reports as Data*. Cambridge, MA: MIT Press, revised edn.

Erkizia, Xabier, Oier Iruretagoiena, Enrike Hurtado, and Xavier Balderas. 2009. Soinu mapa [web-map application]. Available at www.soinumapa.net/?lang=en. Last accessed June 2010.

Faulkner, Laura. 2003. Beyond the five-user assumption: Benefits of increased sample sizes in usability testing. *Behaviour Research Methods, Instruments, and Computers* 35(3):379–383.

Feld, Steven. 1982. *Sound and Sentiment: Birds, Weeping, Poetics and Song in Kaluli Expression*. Conduct and Communication Series. Philadelphia: University of Pennsylvania Press, 2nd edn.

Figgis, Mike. 2003. Silence: The absence of sound. In Sider et al. (2003), chap. 1, 1–14.

Fisher, Peter F. 1994. Hearing the reliability in classified remotely sensed images. *Cartography and Geographic Information Systems* 21(1):31–36.

Fiske, Harold E. 1990. *Music and Mind: Philosophical Essays on the Cognition and Meaning of Music*. Lewiston, New York; Queenston, Ontario; Lampeter, Wales: Edwin Mellen.

Flowers, John H. 2005. Thirteen years of reflection on auditory graphing: Promises, pitfalls, and potential new directions. In *Proceedings of the Eleventh Meeting of the International Conference on Auditory Display (ICAD'05)*. Limerick, Ireland: ICAD. Available at http://digitalcommons.unl.edu/psychfacpub/430. Accessed July 2011.

Flowers, John H., Dion C. Buhman, and Kimberly D. Turnage. 1997. Cross-modal equivalence of visual and auditory scatterplots for exploring bivariate data samples. *Human Factors* 39(3):341–351.

———. 2005. Data sonification from the desktop: Should sound be part of standard data analysis software? *ACM Transactions on Applied Perception* 2(4):467–472. Available at http://digitalcommons.unl.edu/psychfacpub/439/. Accessed July 2011.

Flowers, John H. and Terry A. Hauer. 1995. Musical versus visual graphs: Cross-modal equivalence in perception of time series data. *Human Factors* 37(3):553–569.

Flowers, John H., Laura E. Whitwer, Douglas C. Grafel, and Cheryl A. Kotan. 2001. Sonification of daily weather records: Issues of perception, attention and memory in design choices. In *Proceedings of the 2001 International Conference on Auditory Display*, 222–226. Espoo, Finland: International Community for Auditory Display. Available at http://digitalcommons.unl.edu/psychfacpub/432/. Accessed July 2011.

Francis, Ken. 1996. *Wula Na Lnuwe'katiyek - an electronic mixed media atlas.* Master's thesis, Carleton University, Ottawa.

Freesound Project. 2005. Map of geotagged sounds [web-map application]. Available at www.freesound.org/geotagsView.php. Last accessed June 2010.

———. 2010. Sons de Barcelona [web-map application]. Available at http://barcelona.freesound.org. Last accessed June 2010.

Gartner, Georg. 2004. Location-based mobile pedestrian navigation services: The role of multimedia cartography. In *International Joint Workshop on Ubiquitous, Pervasive and Internet Mapping (UPIMap 2004).* Available at http://www.ubimap.net/upimap2004. Accessed November 22, 2007.

———. 2005. TeleCartography: A new means of geocommunication. In Taylor (2005b), chap. 16, 373–387.

Gaver, William W. 1994. Using and creating auditory icons. In Kramer (1994a), chap. 17, 417–446.

Gaver, William W., Randall B. Smith, and Tim O'Shea. 1991. Effective sounds in complex systems: The ARKola simulation. In *Proceedings of CHI'91*, eds. S. Robertson, G. Olson, and J. Olson, 85–90. Association for Computing Machinery, New Orleans: ACM Press, Addison-Wesley.

Genette, Gérard. 1988. *Narrative Discourse Revisited.* Ithaca, NY: Cornell University Press. Translated by Jane E. Lewin.

Gersmehl, Philip J. 1990. Choosing tools: Nine metaphors of four-dimensional cartography. *Cartographic Perspectives* 5:3–17.

Gitelman, Lisa. 2006. *Always Already New: Media, History and the Data of Culture.* Cambridge, MA: MIT Press.

Gold, John Robert. 1980. *An Introduction to Behavioural Geography.* Oxford; New York: Oxford University Press.

Goldstein, E. Bruce. 1999. *Sensation and Perception.* Pacific Grove, California: Brooks/Cole, 5th edn.

Golledge, Reginald G. 1991. Tactual strip maps as navigational aids. *Journal of Visual Impairment and Blindness* 85:296–301.

Gomery, Douglas. 1976. The coming of sound: Technological change in the American film industry. In Weis and Belton (1985), 5–24.

———. 2005. *The Coming of Sound: A History.* New York: Routledge.

Goodchild, Michael F. 2007. Citizens as voluntary sensors: Spatial data infrastructure in the world of web 2.0. *International Journal of Spatial Data Infrastructures Research* 2:24–32.

Gorbman, Claudia. 1980. Narrative film music. *Yale French Studies* 60 (Cinema/Sound):183–203.

———. 2003. Why music? the sound film and its spectator. In *Movie Music, the Film Reader*, ed. Kay Dickinson, chap. 3, 37–47. London; New York: Routledge.

Gregory, Derek. 2001. (Post)colonialism and the production of nature. In *Social Nature: Theory, Practice, and Politics*, eds. Noel Castree and Bruce Braun, chap. 5, 84–111. London: Routledge.

Guerraz, Agnès and Jacques Lemordant. 2008. Indeterminate adaptive digital audio for games on mobiles. In Collins (2008a), chap. 4, 55–72.

Gunning, Tom. 1986. The cinema of attraction: Early film, its spectators, and the avant-garde. In *Film and Theory: An Anthology*, eds. Robert Stam and Toby Miller, 229–235. Malden, MA; Oxford: Blackwell.

Gwich'in Social and Cultural Institute. 2003. Talking map research. World Wide Web site. http://www.gwichin.ca/Research/placeNameMap.html. Accessed June 2011.

Haggett, Peter. 1965. *Locational Analysis in Human Geography.* London: Edward Arnold.

Hamidi, Foad and Bill Kapralos. 2009. A review of spatial sound for virtual environments and games with graphics processing units. *The Open Virtual Reality Journal* 1:8–17.

Hansen, Mark B. N. 2004. *New Philosophy for New Media*. Cambridge, MA: MIT Press, paperback edn.

Harley, J. Brian. 1989. Deconstructing the map. *Cartographica* 26(2):1–20.

Harley, J. Brian and David Woodward. 1987. *Cartography in prehistoric, ancient, and medieval Europe and the Mediterranean*, vol. 1 of *The History of Cartography*. Chicago: University of Chicago Press.

Harmon, Katherine, ed. 2003. *You are Here: Personal Geographies and Other Maps of the Imagination*, ed. Katherine Harmon. New York: Princeton Architectural Press.

Harrower, Mark. 2003. Tips for designing effective animated maps. *Cartographic Perspectives* 44:63–65.

———. 2007. The cognitive limits of animated maps. *Cartographica* 42(4):349–357.

Hayes, Amos. 2006. Nunaliit: Cybercartographic atlas framework. World Wide Web site published by the Geomatics and Cartographic Research Centre, Carleton University. http://nunaliit.org. Accessed April 2008.

Head, C. Grant. 1991. Mapping as language or semiotic system: Review and comment. In *Cognitive and Linguistic Aspects of Geographic Space*, eds. David M. Mark and Andrew U. Frank, 237–262. Dordrecht; Boston; London: Kluwer Academic Publishers.

Heckroth, Jim. 1996. *The Complete MIDI 1.0 Detailed Specification*. MIDI Manufacturers Association (http://www.midi.org).

Heller, Rachelle S. and C. Dianne Martin. 1995. A media taxonomy. *IEEE Multimedia* 2(4):36–45.

———. 1999. Multimedia taxonomy for design and evaluation. In *Handbook of Multimedia Computing*, ed. Borko Furht, chap. 1, 3–16. Boca Raton: CRC Press.

Hermann, Thomas, Jan M. Drees, and Helge Ritter. 2003. Broadcasting auditory weather reports — a pilot project. In *Proceedings of the 2003 International Conference on Auditory Display*, 208–211. Boston, Massachusetts: International Community for Auditory Display.

Hershenson, Maurice. 1998. *Visual Space Perception: A Primer*. Cambridge, MA: MIT Press.

Hirt, Edward R., Hugh E. McDonald, R. Jeffrey Melton, and Judith M. Harackiewicz. 1996. Processing goals, task interest, and the mood-performance relationship: A mediational analysis. *Journal of Personality and Social Psychology* 71(2):245–261.

Holman, Tomlinson. 2002. *Sound for Film and Television*. Woburn, MA: Focal Press, 2nd edn.

Holzer, Derek, Sara Kolster, and Marc Boon. 2004. Sound transit [web-map application]. Available at http://soundtransit.nl. Last accessed June 2010.

Horrigan, John B. 2006. Online news: For many home broadband users, the Internet is a primary news source. Pew Internet and American Life Project. Available at http://www.pewinternet.org/Reports/2006/Online-News-For-many-home-broadband-users-the-internet-is-a-primary-news-source.aspx. Last accessed February 11, 2010.

Howes, David. 1991a. Introduction: 'to summon all the senses'. In Howes (1991b), 3–21.

Howes, David, ed. 1991b. *The Varieties of Sensory Experience: A Sourcebook in the Anthropology of the Senses*. Toronto: University of Toronto Press.

———. 2005. *Empire of the Senses: The Sensual Culture Reader*. Oxford; New York: Berg.

Hu, Shunfu. 2003. Web-based multimedia GIS: Exploring interactive maps and associated multimedia information on the Internet. In Peterson (2003), chap. 20, 335–344.

Hull, John. 1990. *Touching the Rock: An Experience of Blindness*. London: SPCK Publishing.

Humphrey, David, Corban Brook, Al MacDonald, Yury Delendik, Ricard Marxer, and Charles Cliffe. 2009. Defining an enhanced API for audio. Draft Recommendation. Available at https://wiki.mozilla.org/Audio_Data_API. Accessed August 2011.

IFPI. 2010. IFPI digital music report 2010. Available at http://www.ifpi.org/content/library/DMR2010.pdf. Last accessed February 10, 2010.

ITU-BDT. 2009a. Information society statistical profiles 2009: Africa. Telecommunication Development Bureau, International Telecomunication Union. Available at http://www.itu.int/ITU-D/ict/. Accessed July 2011.

———. 2009b. Information society statistical profiles 2009: Americas. Telecommunication Development Bureau, International Telecomunication Union. Available at http://www.itu.int/ITU-D/ict/. Accessed July 2011.

———. 2009c. Information society statistical profiles 2009: Arab States. Telecommunication Development Bureau, International Telecomunication Union. Available at http://www.itu.int/ITU-D/ict/. Accessed July 2011.

———. 2009d. Information society statistical profiles 2009: Asia and the Pacific. Telecommunication Development Bureau, International Telecomunication Union. Available at http://www.itu.int/ITU-D/ict/. Accessed July 2011.

———. 2009e. Information society statistical profiles 2009: CIS (Commonwealth Independent States). Telecommunication Development Bureau, International Telecomunication Union. Available at http://www.itu.int/ITU-D/ict/. Accessed July 2011.

———. 2009f. Information society statistical profiles 2009: Europe. Telecommunication Development Bureau, International Telecomunication Union. Available at http://www.itu.int/ITU-D/ict/. Accessed July 2011.

Jacob, Christian. 2006. *The Soveriegn Map: Theoretical Approaches in Cartography throughout History*. Chicago; London: University of Chicago Press. Translated by Tom Conley. Edited by Edward H. Dahl.

Jacobsen, Niels Ebbe, Morten Hertzum, and Bonnie E. John. 1998. The evaluator effect in usability tests. In *CHI 98 Conference Summary on Human Factors in Computing Systems*, CHI '98, 255–256. New York, NY: ACM.

Jacobson, R. Daniel. 1994. Navigation for the visually handicapped: Going beyond tactile cartography. *Swansea Geographer* 31:53–59.

Jay, Martin. 1988. Scopic regimes of modernity. In *Vision and Visuality*, ed. Hal Foster, 3–23. Seattle: Bay Press.

Jenkins, Henry. 2006. *Convergence Culture: Where Old and New Media Collide*. New York; London: New York University Press.

Kaae, Jesper. 2008. Theoretical approaches to composing dynamic music for video games. In Collins (2008a), chap. 5, 75–92.

Kalyuga, Slava, Paul Chandler, and John Sweller. 2004. When redundant on-screen text in multimedia technical instruction can interfere with learning. *Human Factors* 46(3):567–581.

Kärjä, Antti-Ville. 2008. Marketing music through computer games: the case of Poets of the Fall and *Max Payne 2*. In Collins (2008a), chap. 2, 27–44.

Karlin, Fred. 1994. *Listening to Movies: The Film Lover's Guide to Film Music.* New York: Schirmer Books.

Karlin, Fred and Rayburn Wright. 2004. *On the Track: A Guide to Contemporary Film Scoring.* New York: Routledge, 2nd edn.

Kassabian, Anahid. 2008. Innattentive engagements: The new problematics of sound and music. *Cinema Journal* 48(1).

Kinayoglu, Gokce. 2010. Soundtracks [web-map application]. Available at www.stoparchitects.com/terrasound/soundtrack. Last accessed June 2010.

Klimek, Mary Pat. 1992. Imagining the sound(s) of Shakespeare. In Altman (1992d), chap. 11, 204–216.

Kozloff, Sarah. 1988. *Invisible Storytellers: Voice-over Narration in American Fiction Film.* Berkely; Los Angeles; London: University of California Press.

———. 2000. *Overhearing Film Dialogue.* Berkely; Los Angeles; New York: University of California Press.

Kraak, Menno-Jan and Allan Brown, eds. 2000. *Web Cartography: Developments and Prospects.* New York: Taylor and Francis.

Kraak, Menno-Jan, Robert Edsall, and Alan M. MacEachren. 1997. Cartographic animation and legends for temporal maps: Exploration and/or interaction. In *Proceedsing of the 18th ICA/ACI International Conference*, 263–260. International Cartographic Association.

Kraak, Menno-Jan and Arjen Klomp. 1995. A classification of cartographic animations: Towards a tool for the design of dynamic maps in a GIS environment. In *Proceedings of the Seminar on Teaching Animated Cartography*, eds. Ferjan Ormeling, Barend Köbben, and Rufino Perez Gomez, 29–37. International Cartographic Association.

Kramer, Gregory, ed. 1994a. *Auditory Display: Sonification, Audification, and Auditory Interfaces.* Santa Fe Institute Studies in the Sciences of Complexity. Reading, MA: Addison-Wesley.

Kramer, Gregory. 1994b. An introduction to auditory display. In Kramer (1994a), chap. 1, 1–78.

———. 1994c. Some organizing principles for representing data with sound. In Kramer (1994a), chap. 7, 185–221.

Kramer, Gregory, Bruce Walker, Terri Bonebright, Perry Cook, John H. Flowers, Nadine Miner, John Neuhoff, Robin Bargar, Stephen Barrass, Jonathan Berger, Grigori Evreinov, W. Tecumseh Fitch, Matti Gröhn, Steve Handel, Hans Kaper, Haim Levkowitz, Suresh Lodha, Barbara Shinn-Cunningham, Mary Simoni, and Sever Tipei. 1997. Sonification report: Status of the field and research agenda. Prepared for the National Science Foundation by members of the International Community for Auditory Display. Available at http://digitalcommons.unl.edu/psychfacpub/444/. Accessed July 2011.

Kramers, R. Eric. 2009. User surveys - "Who, What, Where, When, Why". In *Proceedings of the 24th International Cartographic Conference*, 12 p. Santiago, Chile: International Cartographic Association. Available at http://icaci.org/documents/ICC_proceedings/ICC2009/. Accessed February 2011.

Krygier, John B. 1994. Sound and geographic visualization. In MacEachren and Taylor (1994), chap. 8, 149–166.

Krygier, John B., Catherine Reeves, David DiBiase, and Jason Cupp. 1997. Design, implementation, and evaluation of multimedia resources for geography and earth science education. *Journal of Geography in Higher Education* 21(1):17–39.

Kwan, Mei-Po. 2002. Feminist visualization: Re-envisioning GIS as a method in feminist geographic research. *Annals of the Association of American Geographers* 92(4):645–661.

Lastra, James. 1992. Reading, writing, and representing sound. In Altman (1992d), chap. 3, 65–86.

Lenhart, Amanda, Sydney Jones, and Alexandra Rankin Macgill. 2008. Adults and video games. Pew Internet Project Data Memo. Available at http://www.pewinternet.org/Reports/2008/Adults-and-Video-Games.aspx. Last Accessed February 10, 2010.

Levin, Tom. 1984. The acoustic dimension: Notes on cinema sound. *Screen* 25(3):55–68.

Levine, Paula. 2006. Signature [art installation]. Available at http://paulalevine.net/projects/signature/signature.html. Last accessed June 2010.

Levine, Paula, Geoff Lilleman, Ryan Johnston, Glen Redpath, and Sherrie Dandan. 2004. Baghdad San Francisco [art installation incorporating web application]. Available at http://shadowsfromanotherplace.net. Last accessed June 2010.

Levitin, Daniel J. 2006. *This is Your Brain on Music: The Science of a Human Obsession*. New York: Dutton.

Lewis, Clayton and John Rieman. 1994. Task-centered user interface design: A practical introduction. Available by anonymous download from ftp.cs.colorado.edu.

Lewis, James R. 1994. Sample sizes for usability studies: Additional considerations. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 36:368–378.

Leyshon, Andrew, David Matless, and George Revill. 1998. Introduction: Music, space, and the production of place. In *The Place of Music*, eds. Andrew Leyshon, David Matless, and George Revill, 1–30. New York; London: Guildford.

Lindgaard, Gitte and Jarinee Chattratichart. 2007. Usability testing: What have we overlooked? In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1415–1424. New York, NY.

LoBrutto, Vincent. 1994. *Sound-on-Film: Interviews with Creators of Film Sound.* Westport, CT: Praeger Publishers.

locusonus.org. 2006. Locus sonus [web-map application]. Available at http://locusonus.org/soundmap. Last accessed June 2010.

Lothe, Jakob. 2000. *Narrative in Fiction and Film: An Introduction.* Oxford; New York: Oxford University Press.

MacEachren, Alan. 1994. Visualization in modern cartography. In MacEachren and Taylor (1994), chap. 1, 1–12.

MacEachren, Alan and D. R. Fraser Taylor, eds. 1994. *Visualization in Modern Cartography.* New York: Pergamon.

Macefield, Ritch. 2009. How to specify the participant group size for usability studies: A practitioner's guide. *Journal of Usability Studies* 5(1):34–45.

Madden, Mary. 2009. The audience for online video-sharing sites shoots up. Available at http://fe01.pewinternet.org/Reports/2009/13--The-Audience-for-Online-VideoSharing-Sites-Shoots-Up.aspx. Last accessed February 10, 2010.

Madrid Soundscape.org. 2008. Madrid soundscape map [web-map application]. Available at www.madridsoundscape.org. Last accessed June 2010.

Manovich, Lev. 2001. *The language of new media.* Cambridge, MA: MIT Press.

Manvell, Roger, John Huntley, Richard Arnell, and Peter Day. 1975. *The Technique of Film Music.* London; New York: Focal Press.

Marks, Aaron. 2001. *The Complete Guide to Game Audio: For Composers, Musicians, Sound Designers, and Game Developers*. Lawrence, KA: CMP Books.

Mason, Jennifer. 2002. *Qualitative Researching*. London: Sage, 2nd edn.

Matless, David. 2003. Gestures around the visual. *Antipode* 35(2):222–226.

Maxfield, J. P. 1930. Acoustic control of recording for talking motion pictures. *Journal of the Society of Motion Picture Engineers* 14(1):85–95.

Mayer, Richard E. 2005. Multimedia learning: Guiding visuospatial thinking with instructional animation. In Shah and Miyake (2005), chap. 12, 477–508.

Mayer, Richard E. and Richard B. Anderson. 1991. Animations need narrations: An experimental test of a dual-coding hypothesis. *Journal of Educational Psychology* 83(4):484–490.

———. 1992. The instructive animation: Helping students build connections between words and pictures in multimedia learning. *Journal of Educational Psychology* 84(4):444–452.

McClary, Susan. 1991. *Feminine Endings: Music, Gender, and Sexuality*. Minneapolis, MN; London: University of Minnesota Press.

McGonigal, David and Lynn Woodworth. 2001. *Antarctica and the Arctic: The Complete Encyclopedia*. Willowdale, Ontario: Firefly Books.

McLuhan, Marshall. 1962. *The Gutenberg Galaxy: The Making of Typographic Man*. Toronto: University of Toronto Press.

———. 1964. *Understanding Media: The Extensions of Man*. New York; Toronto; London: McGraw-Hill.

McLuhan, Marshall and Eric McLuhan. 1988. *Laws of Media: The New Science*. Toronto: University of Toronto Press.

Meijer, Peter B. L. 1992. An experimental system for auditory image representations. *IEEE Transactions on Biomedical Engineering* 39(2):112–121. Also see http://www.seeingwithsound.com/voice.html. Last accessed March 31, 2005.

Metzner, Jim and Eileen McAdam. 2010. Sound and story project of the Hudson Valley [web-map application]. Available at www.soundandstory.org. Last accessed June 2010.

Miller, Christopher C. 2006. A beast in the field: The Google maps mashup as GIS/2. *Cartographica* 41(3):187–199.

Miller, Harvey J. 2007. Geographic data mining and knowledge discovery. In *The Handbook of Geographic Information Science*, eds. John P. Wilson and A. Stewart Fotheringham, chap. 19, 352–366. Malden, MA: Blackwell.

Miller, Harvey J and Jiawei Han. 2009. Geographic data mining and knowledge discovery: An overview. In *Geographic Data Mining and Knowledge Discovery*, eds. Harvey J Miller and Jiawei Han, chap. 1, 1–26. Boca Raton, FL: CRC Press, 2nd edn.

Moellering, Harold. 1980. The real-time animation of three-dimensional maps. *The American Cartographer* 7:67–75.

Molina, Berio, Chiu Longina, Horacio González, Juan-Gil Rodríguez, Julio Gómez, and Carlos Suárez Sánchez. c2006. Escoitar [web-map application]. Available at www.escoitar.org/?lang=en. Last accessed June 2010.

Monda, Robin Locke. 2009. Sounds like Staten Island [web-map application]. Available at www.soundslikestatenisland.com/map/node. Last accessed June 2010.

Monmonier, Mark. 1989. Geographic brushing: Enhancing exploratory analysis of the scatterplot matrix. *Geographical Analysis* 21(1):81–84.

———. 1990. Strategies for the visualization of time-series data. *Cartographica* 27(1):30–45.

———. 1992. Authoring graphic scripts: Experiences and principles. *Cartography and Geographic Information Systems* 19(4):247–260, 272.

———. 1994. Graphic narratives for analyzing environmental risks. In MacEachren and Taylor (1994), chap. 11, 201–213.

More, Thomas. 1518. *De Optimo Reipublicae Statu, Deque Nova Insula Utopia*. Basil: Froben, 3rd edn.

Mott, Robert L. 1993. *Radio Sound Effects: Who did it, and How, in the Era of Live Broadcasting*. Jefferson, NC: McFarland.

Mouafo, Dieudonné and Anita Müller. 2002. Web-based multimedia cartography applied to the historical evolution of Iqaluit, Nunavut. In *Proceedings of the Joint International Symposium on Geospatial Theory, Processing and Applications*. International Society for Photogrammetry and Remote Sensing, Commision IV, Spatial Data Handling.

Müller, Jean-Claude and Holger Scharlach. 2001. Noise abatement planning: Using animated maps and sound to visualise traffic flows and noise pollution. In *Proceedings of the 20th International Cartographic Conference*, vol. I, 375–385. International Cartographic Association.

Murch, Walter. 2003. Touch of silence. In Sider et al. (2003), chap. 8, 83–102.

———. 2005. Dense clarity — clear density. *The Transom Review* 5(1):pt. 2. Available at http://transom.org/?p=6992. Last accessed July 2010.

murmurtoronto.ca. 2003. murmur [web-map application]. Available at http://murmurtoronto.ca. Last accessed June 2010.

Murray, Janet Horowitz. 1997. *Hamlet on the Holodeck: The Future of Narrative in Cyberspace*. New York: Free Press.

Mydlarz, Charlie, Trevor Cox, and Ian Drumm. 2009. Sound around you [web-map application]. Available at www.soundaroundyou.com. Last accessed June 2010.

Nichols, Bill. 1985. The voice of documentary. In *Movies and Methods: An Anthology*, ed. Bill Nichols, vol. II, 258–273. Berkeley: University of California Press.

———. 1991. *Representing Reality: Issues and Concepts in Documentary*. Bloomington; Indianapolis: Indiana University Press.

Nielsen, Jakob. 1993. *Usability Engineering*. Boston, MA: Academic Press.

Nielsen, Jakob and Thomas K. Landauer. 1993. A mathematical model of the finding of usability problems. In *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems*, 206–213. New York, NY.

Noll, Udo. c2006. Radio Aporee [web-map application]. Available at http://aporee.org/maps. Last accessed June 2010.

Norman, Donald A. 2004. *Emotional Design: Why We Love (or Hate) Everyday Things*. New York: Basic Books.

Nöth, Winfried. 1990. *Handbook of Semiotics*. Bloomington: Indiana University Press.

Nunavut Youth Consulting, Community of Arctic Bay, Inuit Heritage Trust, Nunavut Arctic College, and Geomatics and Cartographic Research Centre, Carleton University. 2008. Arctic Bay atlas: Spoken names map. Available at http://arcticbayatlas.ca/names.html. Accessed June 2011.

OGC. 2002. Web map service implementation specification. Adopted Specification OGC 01-068r3, Open GeoSpatial Consortium.

———. 2005. Web feature service implementation specification. Adopted Specification OGC 04-094, Open GeoSpatial Consortium.

Oliver, Richard L. 1993. Cognitive, affective, and attribute bases of the satisfaction response. *Journal of Consumer Research* 20(3):418–430.

Olson, Judy M. 2009. Issues in human subject testing in cartography and GIS. In *Proceedings of the 24th International Cartographic Conference*, 7 p. Santiago, Chile: International Cartographic Association. Available at http://icaci.org/documents/ICC_proceedings/ICC2009/. Accessed February 2011.

Olson, Scott R. 2004. Revolution in Middletown: Knowing, being, and media uncertainty. *International Digital Media and Arts Association Journal* 1(1):57–61.

Ondaatje, Michael. 2002. *The Conversations: Walter Murch and the Art of FIlm Editing*. New York: Alfred A. Knopf.

Oracle. 2000. Java runtime environment. Available at http://www.java.com/en/download/index.jsp.

O'Sullivan, David and David Unwin. 2003. *Geographic Information Analysis*. Hoboken, N.J.: John Wiley & Sons.

Paivio, Allan. 1986. *Mental Representations: A Dual Coding Approach*. New York: Oxford University Press.

Papper, Robert A., Michael E. Holmes, and Mark N. Popovich. 2004. Middletown media studies: Media multitasking and how much people really use the media. *International Digital Media and Arts Association Journal* 1(1):4–56.

Paul, Leonard J. 2008. An introduction to granular synthesis in video games. In Collins (2008a), chap. 9, 135–149.

Persson, Donata, Georg Gartner, and Manfred Buchroithner. 2006. Towards a typology of interactivity functions for visual map exploration. In *Geographic Hypermedia: Concepts and Systems*, eds. Emmanuel Stefanakis, Michael P. Peterson, Costas Armenakis, and Vasilis Delis, chap. 15, 275–292. Berlin; Heidelberg: Springer.

Peterson, Michael P. 1995. *Interactive and Animated Cartography*. Englewood Cliffs, N.J.: Prentice Hall.

———. 1996. Between reality and abstraction: Applications of cartographic animation. Available at http://maps.unomaha.edu/AnimArt/article.html. Accessed November 26, 2007.

———. 1999. Active legends for interactive cartographic animation. *International Journal of Geographical Information Science* 13(4):375–383.

Peterson, Michael P., ed. 2003. *Maps and the Internet*. Oxford: Elsevier Science.

Peterson, Michael P. 2007. The Internet and multimedia cartography. In Cartwright et al. (2007), chap. 3, 35–50.

Peterson, Michael P., ed. 2008. *International Perspectives on Maps and the Internet*. Berlin; Heidelberg; New York: Springer.

Peuquet, Donna J. 2002. *Representations of space and time*. New York: Guilford Press.

Picard, Rosalind W. 1997. *Affective Computing*. Cambridge, MA: MIT Press.

———. 1998. Human-computer coupling. *Proceedings of the IEEE* 86(8):1803–1807.

Pickles, John. 2004. *A History of Spaces: Cartographic reason, mapping and the geo-coded world*. London: Routledge.

Pocock, Douglas. 1989. Sound and the geographer. *Geography* 74(3):193–200.

Poll, Andrea and Sha Sha Feng. 2009. Sound seeker [web-map application]. Available at fm.hunter.cuny.edu/nysae/nysoundmap/soundseeker.html. Last accessed June 2010.

Porteous, J. Douglas and Jane F. Mastin. 1985. Soundscape. *Journal of Architectural and Planning Research* 2:169–186.

Pratt, Geraldine. 2000. Research performances. *Environment and Planning D: Society and Space* 18(5):639–651.

Pulsifer, Peter, Sébastien Caquard, and D. R. Fraser Taylor. 2007. Toward a new generation of community atlases: The cybercartographic atlas of Antarctica. In Cartwright et al. (2007), chap. 14, 195–216.

Pulsifer, Peter L., Avi Parush, Gitte Lindgaard, and D. R. Fraser Taylor. 2005. The development of the cybercartographic atlas of Antarctica. In Taylor (2005b), chap. 20, 461–490.

Quaiser-Pohl, Claudia and Wolfgang Lehmann. 2002. Girls' spatial abilities: Charting the contributions of experiences and attitudes in different academic groups. *British Journal of Educational Psychology* 72:245–260.

Quinn, Michelle and Dawn Chmielewski. 2008. Top music seller's store has no doors. *Los Angeles Times* April 4, 2008.

Rainie, Lee. 2010. Internet, broadband, and cell phone statistics. Available at http://www.pewinternet.org/~/media//Files/Reports/2010/PIP_December09_update.pdf. Last accessed February 15, 2010.

Rasouli, Maria, Karen Philp, Shamima Khan, Greg Dunn, Gitte Lindgaard, and Avi Parush. 2004. The application of a user-centered design in the development of two web-based educational cybercartographic atlases. Paper read at 2004 Canadian Association of Geographer Ontario Division (CAGONT) Conference. Waterloo, Ontario.

Ratajski, Lech. 1975. Some aspects of the grammar of map language in terms of cartographic communication. Typescript, 18pp.

Rawes, I. M. 2010. London sound survey [web-map application]. Available at www.soundsurvey.org.uk/index.php/survey/soundmaps. Last accessed June 2010.

Reisberg, Daniel and Frederike Heuer. 2005. Visuospatial images. In Shah and Miyake (2005), chap. 2, 35–81.

Rice, Keith and Steven Rice. 2001. Scoring and scripting music for animated maps. Paper read at the 21st Annual Conference of the North American Cartographic Information Society (NACIS).

Rice, Matt, R. Daniel Jacobson, Reginald G. Golledge, and David Jones. 2005. Design considerations for haptic and auditory map interfaces. *Cartography and Geographic Information Science* 32(4):381–391.

Ritts, Max, Greg J. Smith, Hayley Imerman, Nicholas Thompson, Mia Hunt, Anne Ehrlich, Gavin McMurray, and Neil Wiernik. 2010. Toronto sound ecology [web-map application]. Available at http://torontosoundecology.org. Last accessed June 2010.

Roberts, Donald F. and Ulla G. Foehr. 2008. Trends in media use. *The Future of Children* 18(1):11–37.

Roberts, Lissa. 2005. The death of the new chemist: The 'new' chemistry and the transformation of sensuous technology. In Howes (2005), chap. 6, 106–127.

Roberts, Shelley, Aren Hunter, and Avi Parush. 2008. Usability tests for multimedia maps that use sound. Unpublished report.

Robinson, Arthur H. 1952. *The Look of Maps*. Madison: University of Wisconsin Press.

Robison, Kristopher K. and Edward M. Crenshaw. 2010. Reevaluating the global digital divide: Socio-demographic and conflict barriers to the Internet revolution. *Sociological Inquiry* 80(1):34–62.

Rodaway, Paul. 1994. *Sensuous Geographies: Body, Sense and Place*. London: Routledge.

Rogers, Chris. 2010. Web audio API. W3C Audio Group Proposal. Available at https://dvcs.w3.org/hg/audio/raw-file/tip/webaudio/specification.html. Accessed August 2011.

Rose, Gillian. 1996. Geography as a science of observation: The landscape, the gaze and masculinity. In *Human Geography: An Essential Anthology*, eds. John Agnew, David N. Livingstone, and Alisdair Rogers, chap. 20, 341–350. Cambridge, MA: Blackwell.

———. 1997. Situating knowledges: Positionality, reflexivities and other tactics. *Progress in Human Geography* 21(3):305–320.

———. 2003. On the need to ask how, exactly, is geography "visual"? *Antipode* 35(2):212–221.

Ruoff, Jeffrey K. 1992. Conventions of sound in documentary. In Altman (1992d), chap. 12, 217–234.

Sánchez, J., L. Jorquera, E. Muñoz, and E. Valenzuela. 2002. Virtualaurea: Perception through spatialized sound. In *Proceedings of the 4th International Conference on Disability, Virtual Reality and Associated Technologies*, 197–204. Veszprém, Hungary: ICDVRAT/University of Reading.

Sandelowski, Margarete. 1995. Sample size in qualitative research. *Research in Nursing and Health* 18(2):179–183.

Sanger, George Alistair. 2003. *The Fat Man on Game Audio: Tasty Morsels of Sonic Goodness*. Indianapolis, IN: New Riders.

Schafer, R. Murray. 1977. *The Tuning of the World*. New York: Alfred A. Knopf.

Schobesberger, David. 2009. Towards principles for usability evaluation in web mapping: Usability research for cartographic information systems. In *Proceedings of the 24th International Cartographic Conference*, 7 p. Santiago, Chile: International Cartographic Association. Available at http://icaci.org/documents/ICC_proceedings/ICC2009/. Accessed February 2011.

Servigne, Sylvie, Robert Laurini, Myoung-Ah Kang, and Ki Joune Li. 1999. First specifications of an information system for urban soundscape. In *IEEE International Conference on Multimedia Computing and Systems (ICMCS'99)*, vol. II, 262–266.

Shah, Priti and Akira Miyake, eds. 2005. *The Cambridge Handbook of Visuospatial Thinking*. New York: Cambridge University Press.

Shams, Ladan, Yukiyasu Kamitani, and Shinsuke Shimojo. 2004. Modulations of visual perception by sound. In *The Handbook of Multisensory Processes*, eds. Gemma Calvert, Charles Spence, and Barry E. Stein, chap. 2, 27–33. Cambridge, MA: MIT Press.

Shilling, Russel D. and Barbara Shinn-Cunningham. 2002. Virtual auditory displays. In *Handbook of Virtual Environments: Design, Implementation, and Applications*, ed. Kay M. Stanney, Human Factors and Ergonomics, chap. 4, 65–92. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Shinn-Cunningham, Barbara. 2000. Learning reverberation: Considerations for spatial auditory displays. In *Proceedings of the International Conference on Auditory Display (ICAD 2000)*, 126–134. Atlanta, GA: ICAD.

Sider, Larry, Diane Freeman, and Jerry Sider, eds. 2003. *Soundscape: The School of Sound Lectures, 1998–2001*. London; New York: Wallflower Press.

Siekierska, Eva, Richard Labelle, Louis Brunet, Bill McCurdy, Peter Pulsifer, Monika K. Rieger, and Linda O'Neil. 2003. Enhancing spatial learning and mobility training of visually impaired people: A technical paper on the Internet-based tactile and audio-tactile mapping. *The Canadian Geographer* 47(4):480–493.

Sonnenschein, David. 2001. *Sound design: The Expressive Power of Music, Voice and Sound effects in Cinema*. Studio City, CA: Michael Wiese Productions.

Soskice, Janet. 1996. Sight and vision in Medieval Christian thought. In *Vision in Context: Historical and Contemporary Perspectives on Sight*, eds. Teresa Brennan and Martin Jay, 29–43. New York: Routledge.

Soundwalks.org. 2010. Soundwalks [web-map application]. Available at www.soundwalks.org/map. Last accessed June 2010.

Stanza. 2010. Soundcities [web-map application]. Available at http://soundcities.com. Last accessed June 2010.

Statistics Canada. 2005. Household Internet use at home by Internet activity. CANSIM (database) table 358-0006. Available at http://www.statcan.gc.ca. Last accessed February 3, 2010.

———. 2006. Household Internet use survey, Internet use by location of use and household income, annually (percentage of households). CANSIM (database) table 358-0017. Available at http://www.statcan.gc.ca. Last accessed February 3, 2010.

———. 2009a. Internet use by individuals, by type of activity. CANSIM (database) table 358-0122. Available at http://www.statcan.gc.ca. Last accessed February 3, 2010.

———. 2009b. Internet use by individuals, by type of activity. CANSIM (database) table 358-0130. Available at http://www.statcan.gc.ca. Last accessed February 3, 2010.

———. 2010. Canadian internet use survey, 2009. *The Daily Statistics Canada* 2–4. Available at http://www.statcan.gc.ca/daily-quotidien/100510/dq100510a-eng.htm. Accessed July 2011.

———. 2011. General social survey: Time use. *The Daily Statistics Canada* 6–7. Available at http://www.statcan.gc.ca/daily-quotidien/110712/dq110712-eng.pdf. Accessed July 2011.

Stein, Max. 2008. Montreal sound map [web-map application]. Available at http://cessa.music.concordia.ca/soundmap/en. Last accessed June 2010.

Sterne, Jonathan. 2003. *The Audible Past: Cultural Origins of Sound Reproduction.* Durham, NC; London: Duke University Press.

Stewart, James G. 1980. The evolution of cinematic sound: A personal report. In *Sound and the Cinema: The coming of Sound to American Film,* ed. Evan William Cameron, 38–67. Pleasantville, N.Y.: Redgrave.

Suchan, Trudy A. and Cynthia A. Brewer. 2000. Qualitative methods for research on mapmaking and map use. *Professional Geographer* 52(1):145–154.

Sui, Daniel Z. 2000. Visuality, aurality, and shifting metaphors of geographical thought in the late twentieth century. *Annals of the Association of American Geographers* 90(2):322–343.

———. 2008. The wikification of GIS and its consequences: Or Angelina Jolie's new tatto and the future of GIS. *Computers, Environment and Urban Sytems* 32:1–5.

Sui, Daniel Z. and Michael F. Goodchild. 2001. GIS as media? *International Journal of Geographical Information Science* 15(5):387–390.

———. 2003. A tetradic analysis of GIS and society using Mcluhan's law of the media. *The Canadian Geographer* 47(1):5–17.

Synnott, Anthony. 1991. Puzzling over the senses: From Plato to Marx. In Howes (1991b), chap. 5, 61–76.

Tacchi, Jo. 2002. Radio texture: Between self and others. In *The Anthropology of Media: A Reader*, eds. Kelly Askew and Richard R. Wilk, chap. 14, 241–257. Malden, MA; Oxford: Blackwell.

Taylor, Bret. 2005a. Mapping your way: First announcement of google maps. World Wide Web Log (Blog) published by Google Inc. at http://googleblog.blogspot.com/2005/02/mapping-your-way.html. Accessed April 18, 2008.

Taylor, D. R. Fraser. 1987. Cartographic communication on computer screens: The effects of sequential presentation of map information. In *Proceedings of the 13th International Cartographic Conference*, vol. I, 593–611. Morelia, Michoacán Mexico: International Cartographic Association.

———. 2003. The concept of cybercartography. In Peterson (2003), chap. 26, 405–420.

Taylor, D. R. Fraser, ed. 2005b. *Cybercartography: Theory and Practice*. Amsterdam: Elsevier Science.

Taylor, D. R. Fraser. 2005c. The theory and practice of cybercartography: An introduction. In Taylor (2005b), chap. 1, 1–13.

———. 2008. Six key ideas about cybercartography. Unpublished correspondence.

Taylor, D. R. Fraser and Stephanie Pyne. 2010. The history and development of the theory and practice of cybercartography. *International Journal of Digital Earth* 3(1):2–15.

Tessler, Holly. 2008. The new MTV? Electronic Arts and 'playing' music. In Collins (2008a), chap. 1, 13–26.

The Smalls Limited. 2010. The smalls street sounds [web-map application]. Available at www.thesmalls.com/StreetSounds. Last accessed June 2010.

Théberge, Paul. 2005. Sound maps: Music and sound in cybercartography. In Taylor (2005b), chap. 17, 389–410.

———. 2008. Almost silent: The interplay of sound and silence in contemporary cinema and television. In *Lowering the Boom: Critical Studies in Film Sound*, eds. Jay Beck and Tony Grajeda, chap. 3, 51–67. Urbana, IL; Chicago: University of Illinois Press.

Theofanos, Mary Frances and Janice Redish. 2003. Guidelines for accessible and usable web sites: Observing users who work with screen readers. *Interactions* X(6):38–51.

Thirion, Steph. 2007. Cascade on wheels [video capture of audiovisual map]. Trans Parent Web Design. Available at http://www.trsp.net/cow/. Last accessed June 2010.

Thrower, Norman J. W. 1959. Animated cartography. *The Professional Geographer* 11(6):9–12.

———. 1972. *Maps & Man: an examination of cartography in relation to culture and civilization*. Englewood Cliffs, N.J.: Prentice-Hall.

Truax, Barry. 2001. *Acoustic Communication*. Westport, Connecticut: Ablex Publishing, 2nd edn.

Tuan, Yi-Fu. 1974. *Topophilia: Study of Environmental Perception, Attitudes and Values*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

Tversky, Barbara, Julie Bauer Morrison, and Mireille Betrancourt. 2002. Animation: Can it facilitate? *International Journal of Human-Computer Studies* 57:247–262.

Tye, Michael. 1991. *The Imagery Debate*. Cambridge, MA: MIT Press.

van Dijk, Jan and Ken Hacker. 2003. The digital divide as a complex and dynamic phenomenon. *The Information Society* 19:315–326.

van Dijk, Jan A. G. M. 2005. *The Deepening Divide: Inequality in the Information Society*. Thousand Oaks, California: Sage.

van Elzakker, Corné P. J. M. and Karen Wealands. 2007. Use and users of multimedia cartography. In Cartwright et al. (2007), chap. 34, 487–504.

van Leeuwen, Theo. 2005. *Introducing Social Semiotics*. London; New York: Routledge.

Verstraten, Peter. 2009. *Film Narratology*. Toronto; Buffalo; London: University of Toronto Press. Translated by Stefan van der Lecq.

Vetere, Frank and Steve Howard. 2000. Prior knowledge and redundant multimedia. *2000 IEEE International Conference on Multimedia and Expo 2000* 2:605–608.

Virrantaus, Kirsi, David Fairbairn, and Menno-Jan Kraak. 2009. ICA research agenda on cartography and GI science. *The Cartographic Journal* 46(2):63–75.

Virzi, Robert A. 1992. Refining the test phase of usability evaluation: How many subjects is enough? *Human Factors: The Journal of the Human Factors and Ergonomics Society* 34:457–468.

W3C. 2010. *Scalable Vector Graphics (SVG) 1.1 Specification*, 2nd edn. World Wide Web Consortium. Available at: http://www.w3.org/TR/SVG11/. Accessed October 2010.

Walsh, Chris M. 2010. Year-end sales report: Canadian market. Available at: http://www.billboard.biz/bbbiz/content_display/industry/e3ic0da57b53e8b20f676a4aa12d604f237. Last accessed February 10, 2010.

Warf, Barney. 2001. Segueways into cyberspace: Multiple geographies of the digital divide. *Environment and Planning B: Planning and Design* 28:3–19.

Weis, Elisabeth and John Belton, eds. 1985. *Film Sound: Theory and Practice*. New York: Columbia University Press.

Wenzel, Elizabeth M. 1994. Spatial sound and sonification. In Kramer (1994a), chap. 4, 127–150.

Wild Sanctuary Inc. 2009. Wild Sanctuary [web-map application]. Available at http://earth.wildsanctuary.com. Last accessed June 2010.

Williams, Huw and Rebecca Sumner. 2009. Listen to Africa [web-map application]. Available at www.listentoafrica.com. Last accessed June 2010.

Williams, Sheila M. 1994. Perceptual principles in sound grouping. In Kramer (1994a), chap. 3, 95–126.

Wood, Denis. 1987. Pleasure in the idea / the atlas as narrative. *Cartographica* 24(1):24–46.

Wood, Denis and John Fels. 1986. Designs on signs / myth and meaning in maps. *Cartographica* 23(3):54–103.

Woodruff, Andy. 2009. In need of a dancing banana. Published as an entry in the *Cartastrophe* web log. Available at http://cartastrophe.wordpress.com/tag/bad-colors/. Accessed August 2011.

Ylirisku, Salu and Jacob Buur. 2007. *Designing with Video: Focusing the User-Centred Design Process*. London: Springer.

Zamaria, Charles and Fred Fletcher. 2008. *Canada Online! The Internet, media and emerging technologies: Uses, attitudes, trends and international comparisons 2007*. Toronto: Canadian Internet Project. Available at www.ciponline.ca. Last accessed February 10, 2010.

Appendices

Appendix A: Survey of 'Sound Maps' on the World Wide Web: June 2010

Tables A.1 and A.2 present results from a review of web-mapping application designs found by searching the World Wide Web during June 2010. These web-map applications were found by 1) initially querying the Google and Google Scholar search engines using the terms "sound map", "acoustic map", and "audiovisual map"; and 2) by following URLs to "similar" or "highlighted" web applications listed on sites identified in the step 1 searches. See Section 1.3 for a discussion of the results of this survey.

Table A.1: Web-map applications with sound embedded in the map. Results of review of audiovisual web-maps, June 15–22, 2010.

#	Map Title	Attribution	URL
1	BadiaFonia	Cordobés and Picanyol (2008)	www.badiafonia.net
2	Folk Songs for the Five Points	Dant et al. (c2006)	www.tenement.org/folksongs/client
3	Manchester: Peripheral	Dant et al. (2007b)	www.manchesterperipheral.com/phase1/soundmap
4	Radio Aporee	Noll (c2006)	http://aporee.org/maps
5	Sound Map: The Caldeonian Road	Dein and Panetta (2010)	www.guardian.co.uk/society/interactive/2010/apr/26/caledonian-road-sound-map
6	Soundtracks	Kinayoglu (2010)	www.stoparchitects.com/terrasound/soundtrack
7	Soundwalks	Soundwalks.org (2010)	www.soundwalks.org/map

Table A.2: Web-map applications organizing sounds within a map. Results of review of audiovisual web-maps, June 15–22, 2010. Continuing on next page ...

#	Map Title	Attribution	URL
1	12 Gates to the City: Acoustic Map	12 Gates to the City (2009)	http://12gatestothecity.com/extras/acoustic-map
2	Andalucía Soundscape	Cantizzani et al. (c2006)	www.andaluciasoundscape.net/mapa
3	audioBus	Callard et al. (2006)	www.audiobus.org
4	BBC Save Our Sounds	BBC (2009)	www.bbc.co.uk/worldservice/specialreports/saveoursounds/index.shtml
5	Chicago Sound Map	Chicago Public Radio (2007)	www.chicagopublicradio.org/Soundmarks_Browse.aspx
6	Cinco Cidades	Dant et al. (2007a)	www.cincocidades.com/en/soundmap
7	City of Memory	City Lore and Local Projects (2002)	www.cityofmemory.org/map/index.php
8	Escoitar	Molina et al. (c2006)	www.escoitar.org/?lang=en
9	Listen to Africa	Williams and Sumner (2009)	www.listentoafrica.com
10	Listening to Nature: A Sound Walk across California	CLNS and OMCA (c2003)	http://museumca.org/naturalsounds/home.html
11	Locus Sonus	locusonus.org (2006)	http://locusonus.org/soundmap
12	London Sound Survey	Rawes (2010)	www.soundsurvey.org.uk/index.php/survey/soundmaps
13	Madrid Soundscape Map	Madrid Soundscape.org (2008)	www.madridsoundscape.org
14	Map of Geotagged Sounds	Freesound Project (2005)	www.freesound.org/geotagsView.php
15	Memoryscape Audio Walks	Butler and Whiting (2005)	www.memoryscape.org.uk
16	Mississauga Sound Map	Centeno and Sinclair (2008)	www.yorku.ca/caseaces/soundmap
17	Montreal Sound Map	Stein (2008)	http://cessa.music.concordia.ca/soundmap/en
18	murmur	murmurtoronto.ca (2003)	http://murmurtoronto.ca
19	Open Sound New Orleans	Booth and Brancasi (2008)	www.opensoundneworleans.com/core
20	Puget Soundscape	Becker (2005)	http://homepage.mac.com/brettbecker/soundscape/pugetsoundscape.html
21	Soinu Mapa	Erkizia et al. (2009)	www.soinumapa.net/?lang=en
22	Sons de Barcelona	Freesound Project (2010)	http://barcelona.freesound.org
23	Sound and Story Project of the Hudson Valley	Metzner and McAdam (2010)	www.soundandstory.org
24	Sound Around You	Mydlarz et al. (2009)	www.soundaroundyou.com

Table A.2: Web-map applications organizing sounds within a map (continued).

#	Map Title	Attribution	URL
25	Sound Seeker	Poll and Feng (2009)	fm.hunter.cuny.edu/nysae/ nysoundmap/soundseeker.html
26	Sound Tourism: A Travel Guide to Sonic Wonders	Cox (2010)	www.sonicwonders.org
27	Sound Transit	Holzer et al. (2004)	http://soundtransit.nl
28	Soundcities	Stanza (2010)	http://soundcities.com
29	Sounds Like Staten Island	Monda (2009)	www.soundslikestatenisland.com/ map/node
30	Soundscape of China	Eason and Lewis (2007)	www.pbs.org/kqed/chinainside/ soundmap/index.html
31	The Smalls Street Sounds	The Smalls Limited (2010)	www.thesmalls.com/StreetSounds
32	Toronto Sound Ecology	Ritts et al. (2010)	http://torontosoundecology.org
33	Wild Sanctuary	Wild Sanctuary Inc. (2009)	http://earth.wildsanctuary.com

Appendix B: Functional Description of Evaluated Audiovisual Maps

This Appendix describes the map examples evaluated as part of this thesis as discussed in Chapter 6. Videos showing the examples being used are included in the CD-ROM accompanying this thesis (see Appendix G for details).

For the purposes of describing the materials presented to participants during the user evaluation tests, the current discussion is limited to functional descriptions of the interactive audiovisual maps. The focus is *what* the maps do including initial presentation when accessed and in response to the actions of users working with the maps. Chapter 7 provided an overview of the technologies developed as part of this research that assist in the design and implementation of audiovisual maps.

Table B.1 contains a list of the audiovisual map examples assessed during the user evaluations and the Universal Resource Locator (URL) to be used to access each of them on the World Wide Web. Appendix G contains a listing of video captures showing the maps being used as an alternative means of seeing how these map examples function.

B.1 Ottawa Area Federal Election Sound Map: June 28, 2004 (OAFESM)

I designed *Ottawa Area Federal Election Sound Map: June 28, 2004* showing results from the Canadian federal election of that date for a selection of electoral districts in the vicinity of Ottawa, Ontario to demonstrate the use of sound in a map intended for World Wide Web browsing.[1] Figure B.1 shows the web application displaying the election results visually along with a description detailing how a user would work with the map to hear the embedded sounds. The map shows the *first past the post* results of the election with each electoral district shaded only according to the party affiliation of the winning candidate.

This view on its own reduces the election results to only the most basic repre-

[1]Brauen (2006) describes this map design and the technologies used to create it in detail and the functional description of the behaviour of the application presented here is adapted from that article.

Table B.1: Example audiovisual web-map applications assessed during user evaluation testing.

Map Title	URL
Ottawa Area Federal Election Sound Map: June 28, 2004 (OAFESM)†	http://gcrc.carleton.ca/cne/proof_of_concepts/elect2004/
Canada's Trade with World Regions: 1976–2000 (CTWR)†	http://atlas.gcrc.carleton.ca/tdb/world
Antarctic Territorial Claims (ATC)‡	http://atlases.gcrc.carleton.ca/antarctic/territorial/territories.xml.html
Antarctic Exploration (AE)‡	http://atlases.gcrc.carleton.ca/antarctic/exploration/exploration.xml.html

† Design and Production by Glenn Brauen.

‡ Courtesy of Geomatics and Cartographic Research Centre, Carleton University (Caquard et al. 2008; Pulsifer et al. 2007, 2005).

sentation of the outcome. First, all districts won by any given party are represented equivalently according to this scheme.[2] This is appropriate in one sense. Regardless of the percentage of the vote taken by the winning candidate, a single Member of Parliament obtains a seat in Canada's legislature. But the extent of the challenge faced by candidates in distinct districts may have been significantly different for a variety of reasons, including individual candidate popularity or qualitative regional differences. As shown in Table B.2, candidates for the Liberal party were elected in the Ottawa-Vanier and Pontiac districts, although the winning candidate in Ottawa-Vanier received a larger share of the votes (an increase of 11%). In addition, the tabulated result shown for the Bloc Québécois candidate in Pontiac indicates that at least some of the Liberal candidate's opposition in Pontiac was related to Québec regional politics. Neither of these issues is apparent when comparing these districts as represented strictly visually on the map shown in Figure B.1 because it shows only the final result and provides no insight concerning election issues or sources of opposition or support within each district.

In addition to obscuring comparison of election results between districts, the *first past the post* map simplifies the results within a district. Of the electoral districts included in this map, only Argenteuil-Mirabel and Carleton-Lanark elected representatives by giving them a majority of the vote (> 50% + 1). The elected candidates in Ottawa Centre and Pontiac, by contrast, won contests in which 44%

[2]Although geographically large electoral districts can tend to dominate the map in comparison to smaller, more densely populated districts. This is a known problem with categorical thematic maps of varying sized regions and of choropleth maps with varying sized regions or varying population densities.

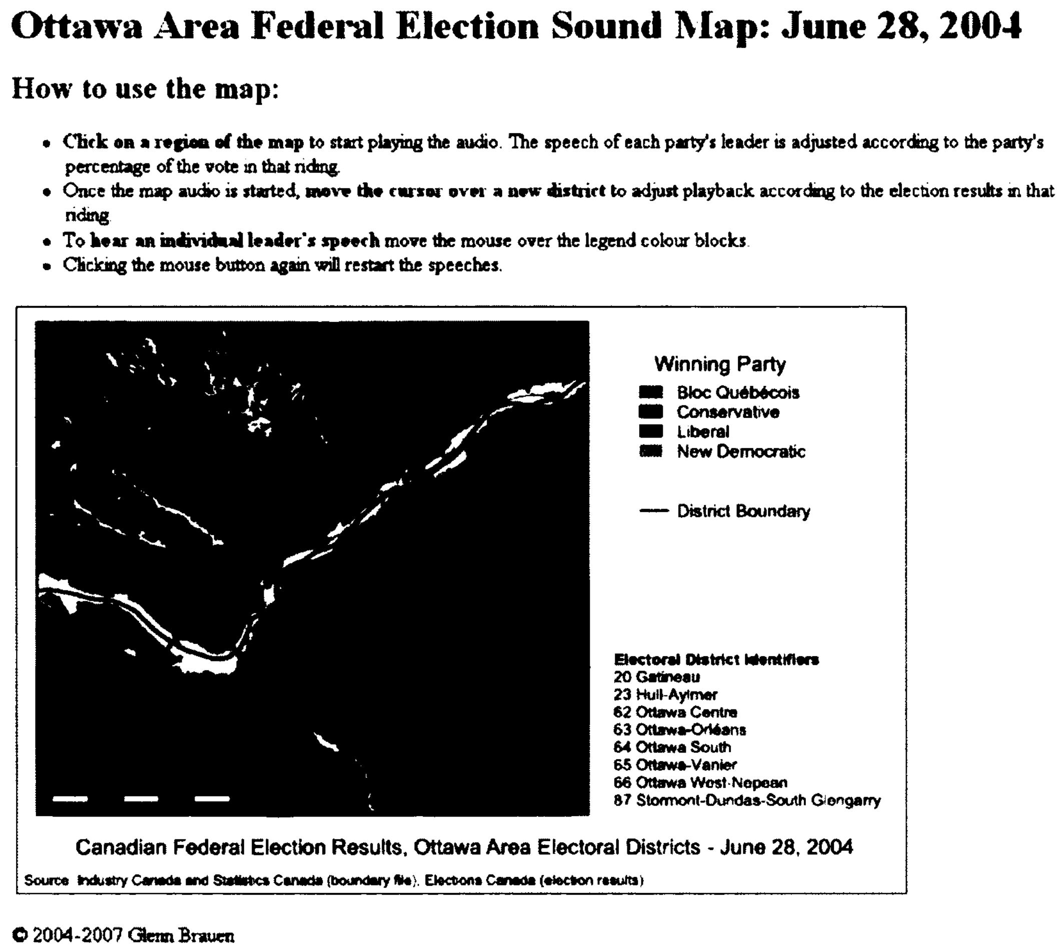


Figure B.1: Ottawa Area Federal Election Sound Map: June 28, 2004 evaluation version web page (application design and production: Glenn Brauen).

and 60% more votes, respectively, were cast against than for them. By reducing the results for each district to a single designation of the winning party, the visual map hides the level of debate (i.e., the presence of other voices) within a district.

To experiment with alternatives that would represent the party affiliation of the winning candidate in each riding while attempting to differentiate those results based on the votes received, I created two variations. The first, beyond the scope of this thesis, modified the *first past the post* map by desaturating the colour used to represent each winning candidate's party based on the final vote percentages in the electoral district. The second, to be discussed here, kept the visual representation as shown in Figure B.1 and augmented it with a sound design intended to: differentiate the results between electoral districts based on vote percentages within each district; represent multiple vote percentages within each district rather than just that of the elected candidate; and add qualitative information concerning is-

Table B.2: Selected Election Results for Ottawa area Districts: June 28, 2004 (Elections Canada 2004).

Party	Results By Selected Riding (%)				
	Argenteuil-Mirabel	Carleton-Lanark	Ottawa Centre	Ottawa-Vanier	Pontiac
Bloc Québécois	57.4	na†	na†	na†	29.2
Conservative	7.0	50.0	19.0	24.2	22.2
Green	5.1	5.6	7.5	6.9	4.2
Liberal	26.9	34.0	31.4	49.2	38.4
New Democratic	3.0	10.4	40.9	18.5	5.8
Independent / Other	0.6	0.0	1.3	1.2	0.3

† Not applicable: Bloc Québécois nominates candidates only within the province of Québec.

sues discussed during the election. When viewed in a web browser, a user may click the mouse while the cursor is over the map, causing recorded speeches of the leaders of the federal parties to play, with the sound level of each leader adjusted according to the vote percentage received by that party's candidate in the riding over which the cursor is positioned. Moving the cursor over another district will cause all of the speech sound levels to be readjusted according to the percentages in the new district. Positioning the cursor over the legend symbols will cause the speech of a single party's leader to be played alone.

The resulting map, while still providing a simple visual impression of the party affiliation of the winning candidate in each district, reintroduces some of the complexity of the election results into the map. Placing the cursor over a district such as Carleton-Lanark in which the winning candidate won a strong majority, the speech of the winning candidate's party's leader is dominant, although other voices are apparent as well. If the cursor is moved over a more closely contested district, such as Ottawa Centre or Pontiac, several leaders' speeches blend together creating a cacophonous result.

The speeches used in the map are cultural objects and do not function solely as formal elements of the map intended to show the relative vote counts for the different parties in each district. If that was the only goal, other methods may have been more effective. As mentioned above, visual mapping techniques could have been used as could visual interactive methods for drilling into the data as the cursor moves over each district (in a manner analogous to the way in which the sound design described is driven through user actions that select electoral districts). Using audiovisual mapping methods, more consistent sounds than speech, containing as it does wide variations in pacing, timbre, tone, and loudness, could have been selected to represent the vote percentage associated with each party. Rather, as Théberge (2005) argued, the sounds chosen are cultural objects on their own and bring additional meanings into the map created using them. By choosing a selection of speeches used by the party leaders during the election campaign, the map makes brief mention of the issues being discussed by the parties during and immediately after the election (the map contains campaign and election night speeches). The result is at once a subjective representation of the election results and, literally, a reintroduction of voices, silenced in the original map, representing the parties of the unsuccessful candidates in each riding.

Both the original and a technologically updated version of *Ottawa Area Federal Election Sound Map: June 28, 2004* are available on the World Wide Web[3] and a video of the map being used is included on the CD-ROM enclosed with this thesis (see Appendix G). For the user evaluation sessions, the visual design of the web page containing the election map was simplified, as shown in Figure B.1, to remove distractions such as technical discussions of software compatibility. The text included on the page describes what a user should do to cause the web page to start producing sounds and states that the party leaders' speeches "are adjusted according to the party's percentage of the vote." The textual description does not describe why sounds were added to the map, nor does it describe why these particular speech fragments were selected.

B.2 Canada's Trade with World Regions: 1976–2000 (CTWR)

Figure B.2 shows partial views of a World Wide Web application prototype, *Canada's Trade with World Regions: 1976–2000* (CTWR), that I designed to complement existing content in the *Cybercartographic Atlas of Canada's Trade with the World* (Eddy and Taylor 2005). Using this an interested user to browse measurements of Canada's trade with world regions using an audiovisual interface. Data observations can be selected representing the economic value of commodity exports and imports between Canada and regions of the world, aggregated geographically at approximately a continental scale and temporally to five-year periods. The data may be examined using top-level commodity sector categories (*Materials and Energy, Food and Agriculture,* and *Manufactured Goods*) or each of these sectors may be refined further.

As a user works with the application (hereafter referred to as *CTWR*), linked visual and auditory representations present import and export measurements. Absolute exchange values may be selected for visual display by selecting a region,

[3]The earliest version of *Ottawa Area Federal Election Sound Map: June 28, 2004* as described in Brauen (2006) is accessible at http://gcrc.carleton.ca/cne/proof_of_concepts/elect2004/ and is identified at that URL as the "original (Adobe) version". This version of the map is effectively obsolete, having been built to use proprietary extensions to the Scalable Vector Graphics standard (SVG, see W3C 2010) for which Adobe has more recently dropped support. A more recent version of the map, identified as the "Java version" at that web site, can be accessed and should work properly if viewed using a web browser that supports standard SVG and for which a Java run-time environment (Oracle 2000) has been installed (more details are provided on the web page).

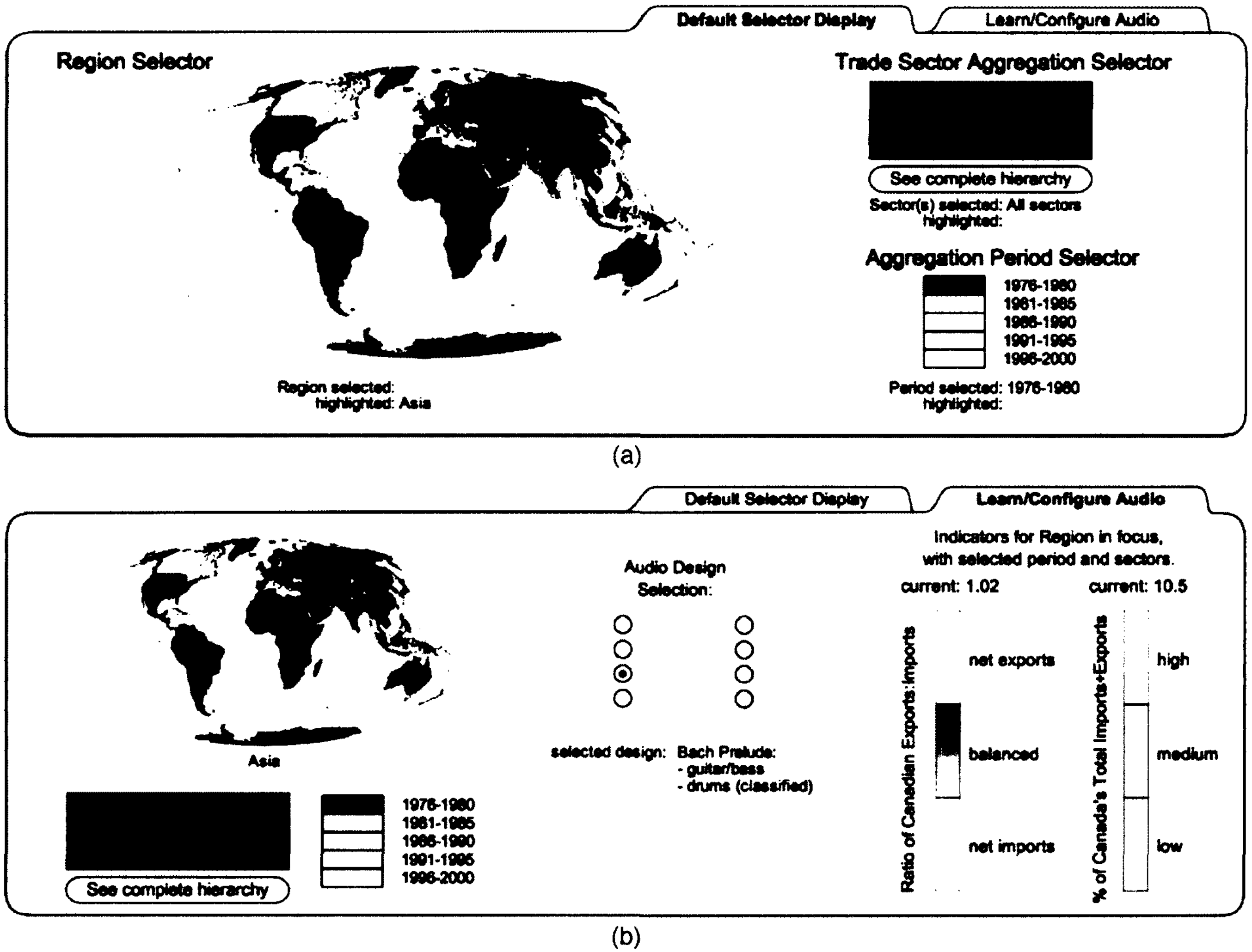


Figure B.2: Canada's Trade with World Regions: 1976–2000, partial views showing a) default display, and b) learn and configure audio display (design and production: Glenn Brauen).

sector aggregation level, and five-year period and requesting that bar charts for the selection be displayed. A pair of charts, one showing exports and the other showing imports are displayed. CTWR currently provides display space, separated from the interactive map of world regions, for up to three bar chart sets to be shown at once and a user may select each of the three according to any criteria.[4]

Relative measurements of trade between regions are represented acoustically using synthesized music with playback controlled by a user's current data selections, consisting of a combination of the choices for trade sector and date range and the current cursor position. As soon as a user moves the cursor over any of the map regions, the auditory representation of these variables will begin and, unlike the version of *Ottawa Area Federal Election Sound Map: June 28, 2004* used during the evaluations, no initial mouse click to start the audio is required. The variables represented using sound are: a) *balance of trade*, computed as the ratio of Canada's exports to the region in comparison to Canada's imports from the region; and b) *value of exchange as a percentage of Canada's total trade* (value of exports + value of imports).

To help a user understand these auditory representations, CTWR provides accompanying visual feedback in response to user actions. To provide regional context for the sounds as a user moves the cursor over the map, it is visually updated to emphasize an outline of the currently highlighted region, as shown in Figure B.2 in which Asia is highlighted, and a label for the region is displayed below the map. Analogous to the use of a visual legend to explain a thematic map representation, *dynamic audiovisual legends* are included as part of the 'Learn/Configure Audio' view of the application (Figure B.2b) to assist a user in understanding the sound design and modulations as a result of variable updates as they work with the map. As shown in Figure B.3a, the dynamic legends act as visual learning aids, providing a correlating visual depiction for the current value of each variable represented acoustically as a user works with the map. In addition, as shown in Figure B.3b, the dynamic legends may also be directly manipulated by a user to adjust the value of a single variable by positioning the cursor at different positions over the visual legend, each position used to derive a value for the associated variable such that repeated auditory updates caused by the cursor moving over the legend can be used to reveal how the auditory representation changes for different values of that variable. Because only the auditory representation for the variable associated with the manipulated legend is heard while the other is silenced, direct manipulation of the legends allows a user to isolate the acoustic representation of

[4]This application is a redesign of an earlier interactive map created for the *Cybercartographic Atlas of Canada's Trade with the World* that placed regional bar charts on a world map. The charts were separated from the map: a) to provide more space for the bar charts than was done in the atlas map; and b) to produce an audiovisual map application that could be assessed as part of user evaluation sessions concerning maps incorporating sound, ensuring that it could be set up for use such that, at least in some cases, participants could only answer questions using information provided aurally.

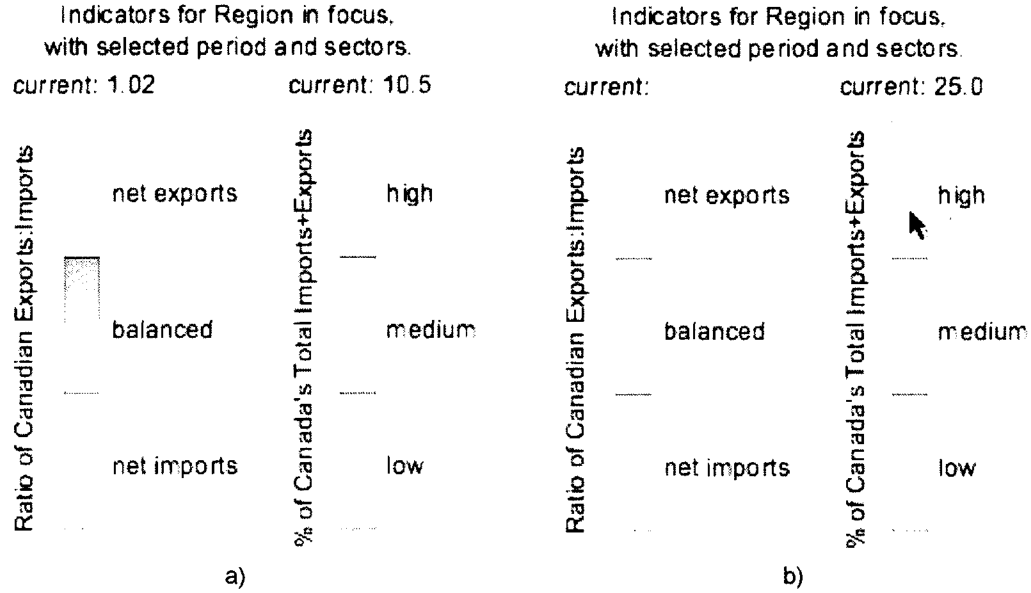


Figure B.3: Dynamic audiovisual legends: a) showing variables for highlighted region, and b) as audio controllers (application design and production: Glenn Brauen).

each variable. As shown, the other dynamic audiovisual legend is also visually cleared. Once the auditory representation has been learned, a user may be able to work with the auditory representation alone and the two-panel design of this application was motivated in part by the desire to test that hypothesis.[5]

As shown in Figure B.2b, a user working with CTWR can select alternative sound designs using a set of radio buttons in the 'Learn/Configure audio' view. Having decided to design the sound using synthesized music, many possibilities spanning music selection, instrument choices, and methods for controlling and modulating the playback of the selected music yet remained. The selectable designs are the results of different choices concerning, among other things, the selection and number of instruments used, choice of metaphors for modifying the playback of sounds based on the current value of an associated variable (e.g., *louder means more*), and whether or not to classify variable values as part of the representation. The current sound designs for CTWR use either synthesized music alone or a combination of synthesized music and recorded audio. Each selection also defines associated control data that specifies the auditory parameters associated with each instrument in the composition or with the recorded sound to be adjusted and how they will be modified in response to variable value changes.

CTWR, as described thus far, is available on the World Wide Web[6] and can be

[5]Abstract uses of sound such as the uses of music described here almost certainly require that a learning aid of some form be available for a user. Alternatively, some coded, semantic uses of sound such as uses of language may be understandable without this type of assistance.

[6]http://atlas.gcrc.carleton.ca/tdb/world

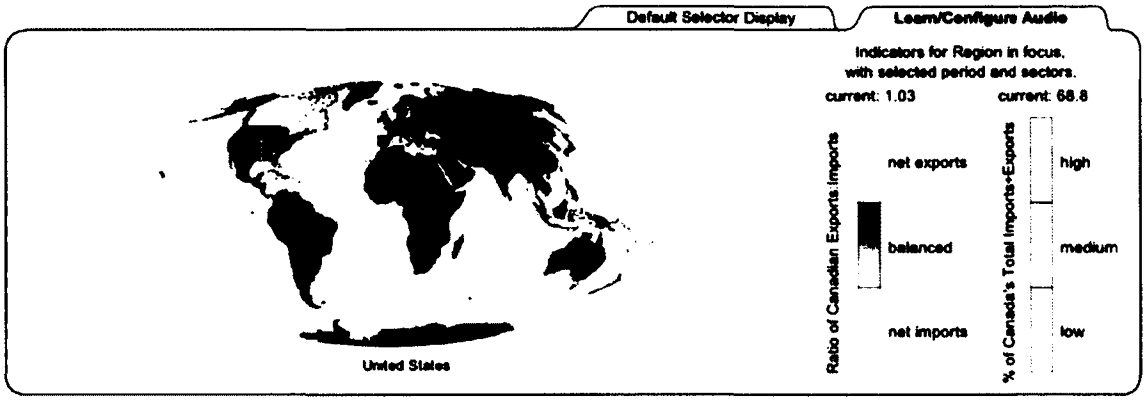


Figure B.4: Canada's Trade with World Regions: 1976–2000, user evaluations version (design and production: Glenn Brauen).

seen in a set of videos showing the map, or various elements of the map, being used that are included on the CD-ROM enclosed with this thesis (see Appendix G). For the user evaluation sessions, CTWR was simplified to focus the use and discussions during the sessions on the use of sound and to avoid having to explain possible non-geographic data selection criteria (time and trade sector) and data semantics. The date range, trade sector, and sound design selectors were removed from the interface and the map was hard-coded to present an overview of trade for all sectors during the period 1976–1980 and using a default sound design. In addition, the display panel for the bar charts was removed so that only the acoustically represented relative measurements (trade balance and value of exchange as a percentage of Canada's total international trade) were available. The resulting version of CTWR retained the two-panel design allowing users to use the dynamic audiovisual legends initially to understand the sound design (Figure B.4) and allowing the interviewers to ask users to work without the visual aid to answer questions later in the session. The alternate panel view for the evaluation version of CTWR contains only the interactive world map and the text label identifying the currently highlighted region. There is no explanatory text in either panel describing how to work with the map, or discussing what the sounds mean or why the map is not silent. The remainder of this section describes the sound design chosen for use during the evaluation sessions.

The version of CTWR used during the evaluation sessions incorporated a sound design in which the variables represented acoustically use classification. This is shown in the visual design of the dynamic audiovisual legends (Figure B.3) and is apparent in the way the acoustic representation for each variable is controlled, with the acoustic representation only being modified if the current value is updated such that the classification of that value changes from the classification of the previous value (i.e., the value crosses a class boundary). In such an implementation, values for each of the control parameters that adjust the playback of the sound representing a spatial variable, and indirectly the mixing of all auditory representations, are assigned to each category of the classification. As regions of the map are selected through interactive control, or when the dynamic legend is used as an audio control, a change in the classification of the associated variable causes the cartographic sound subsystem to be updated with all of the control parameter values associated with the new category in the definition of the variable's acoustic representation.

It is also possible to use auditory representation with unclassified variables (see Chapter 7) and to design the visual legend to reflect this usage. Visually, a dynamic audiovisual legend for a linearly interpolated representation such as that shown in Figure 7.7b should be designed to look like a continuum and a user's interactions with it and the map should cause it to be updated to indicate the current value within that continuum (as opposed to the updates only on class boundary crossings as described above). Nothing restricts the control parameters for a non-classified variable to being computed as shown for this example and it would be valid, in some cases, to consider using logarithmic scales or other non-linear scales as long as the usage could be communicated clearly to the user. Ultimately, design choices will need to be tested to verify that the overall audiovisual design is understood by users.

The *balance of trade* variable, as used during the evaluation sessions, is based on the looping playback of 16 measures from a lute prelude by Johan Sebastian Bach (BWV 997),[7] and controlled according to the current positioning of the cursor over the application's map or the dynamic audiovisual legends as outlined in Table B.3. The MIDI SMF composition, originally a single instrument part, was separated into two instrument parts with the bass harmony played on cello and the melody played on nylon string guitar. The controls specified for the synthesized composition use the cello and guitar parts together to voice balanced trade, metaphorically representing balanced trade by balancing the two instrument parts. Alternatively, variable values classified as *net imports* or *net exports* are represented by one or the

[7]Some characteristics of this composition that led to its use as one of the sound designs in CTWR include: a) temporal consistency in each instrumental part; b) easily distinguishable instrumental parts; and c) the composition and a MIDI SMF encoding of the piece are both available in the public domain. Firstly, the opening 16 measures consist of temporally consistent bass and melody lines (i.e., no rests in either part) which could be separated and assigned to different instruments allowing them to be controlled separately. The lack of rests should help to prevent misinterpretation on the part of a user that, for example, a silence in the guitar part was caused by the current position of the cursor rather than by a pause in the instrumental part. Secondly, the resulting instrumental parts are easy to distinguish, differing as they do along multiple auditory characteristics including timbre, register, and tempo. Thirdly, the composition is in the public domain and a MIDI encoding of a performance of the composition, compatible with the audio subsystem software and music editing software needed to modify the piece for use with the sound subsystem (e.g., separating the single instrument line into the two parts as described) is also in the public domain. The MIDI SMF is *Lute Suite BWV 997: Prelude by J. S. Bach* and is available from the Mutopia Project, http://www.mutopiaproject.org/cgibin/piece-info.cgi?id=50.

Table B.3: Audio components and control parameters used in sample sound design for Canada's trade with the World.

Variable	Instrument	Purpose
trade balance	nylon string guitar	audible indicates either *net exports* if heard alone or *balanced* if heard accompanied by cello.
	cello	audible indicates either *net imports* if heard alone or *balanced* if heard accompanied by nylon string guitar.
value of exchange	drums	loudness is proportional to value of exchange between Canada and the highlighted region; *louder means more*

other of the two parts being muted as detailed by Table B.3.

The combined nylon string guitar and cello representation for balance of trade uses a *'higher pitch and tempo means happier'* metaphor based on a mercantilist understanding of trade relations that suggests that trade surpluses are more positive than trade deficits, with the understanding that many users of this map would be Canadian and thus reflecting that perspective.[8]

An admittedly incongruous synthesized drum track was added to represent the *value of exchange* for the selected region with the loudness of the drum track representing the value of Canada's trade with the currently highlighted region using a *'louder means more'* metaphor. To indicate this, the following relation always hold (at least approximately)[9] between the current loudness of that sound and the value of trade between Canada and the currently selected map region:

$$\frac{\text{Value of Canadian trade with selected region}}{\text{Value of total Canadian trade}} \propto \frac{\text{Current loudness}}{\text{Loudest possible setting}} \tag{B.1}$$

Using classified representation, as is done here, the proportional relation is maintained by defining the loudness control (*gain*) parameter settings such that higher trade value classifications (i.e., *low*, *medium*, *high* in that order) are assigned higher gain values according to a reasonable gradation.

[8] Memory of this metaphor is further supported in the visual design of the dynamic audiovisual legend by placing the 'net exports' class vertically above the 'balanced' and 'net imports' classes. This placement corresponds to the location of a high-register melody line above a low-register bass line on a musical staff. Whether or not a user would trace the metaphor through this number of levels in the visual and auditory design can be validly questioned.

[9] Sound perception, accounting for variations in acoustic parameters (e.g., pitch, timbre), playback system characteristics (e.g., use of speakers versus headphones, let alone the range of qualities in speaker or headphone designs), and the hearing capabilities of listeners all vary greatly. Strictly controlling audio parameters according to an idealized model may seem like an obvious choice but may not always provide greater value than using a reasonable selection of settings that allow you to easily distinguish the different categories as represented and then being prepared to test your settings with users.

B.3 Maps from the Cybercartographic Atlas of Antarctica

The *Cybercartographic Atlas of Antarctica* was developed as a transdisciplinary research project organized by Carleton University's Geomatics and Cartographic Research Centre to which many collaborators and researchers contributed (Pulsifer et al. 2007, 2005). I, as part of a group researching the design and use of sound within the context of cybercartographic atlases such as this one, collaborated on several sections of the atlas, two of which subsequently were included in the user evaluations described here: one concerning national claims to territories of Antarctica and the other concerning a selection of notable explorations around, across, or over the continent. Caquard et al. (2008) discuss the motivations, goals, and technological approaches used in designing sounds for this atlas.

B.3.1 Antarctic Territorial Claims (ATC)

The *Territorial Claims* section of the atlas is composed of subsections that explain the rationale for claims by nations to parts of the territory of Antarctica, the sectoral system of division that has been used to demarcate those claims, the contemporary status of claims to territory under the Antarctic Treaty System, and a summary of those national claims detailing when each was made and the territory claimed by each.

One of these subsections highlights the contestation by the United Kingdom, Argentina, and Chile for overlapping territories of Antarctica. As shown in Figure B.5, the interactive map central to this subsection displays the national claims to territory as semi-transparent sectors overlaying the continent. Each sector is defined as the area between two radii extending from the South Pole to points on the circle of latitude at 60° South.[10] As can be seen, the semi-transparent territorial sectors are not a uniform shade over the entire area of the map. This is due to the presence of overlapping claims on the map and the consequent layering of semi-transparent sectors, complicating the otherwise simple seeming sectoral division of the continent.

Sound is used to reinforce the complexity of the overlapping territorial claims in the visual map, to suggest the tension inherent in the existence of overlapping claims, and to provide contextual information about each of the national claims. For each of the nations with overlapping claims, a recorded narration of a newspaper story discussing Antarctica from a perspective aligning with that national claim and read in the official language of that claiming nation (i.e., English for the

[10]Rather than labelling each of the sectors on the map, each sector and the names of the claiming nations are hyperlinked within the web application so that positioning the cursor over the text for a nation's name causes the associated map sector to be highlighted (and vice versa).

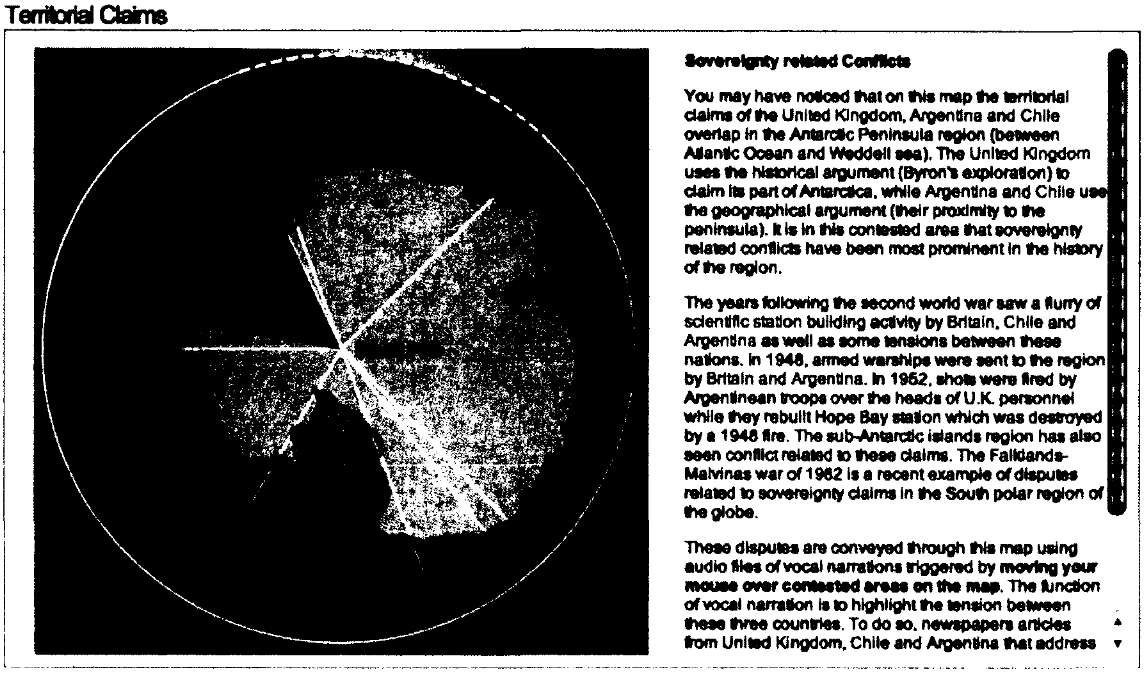


Figure B.5: Cybercartographic Atlas of Antarctica: Territorial Claims (image courtesy of the Geomatics and Cartographic Research Centre, Carleton University).

U. K. and Spanish for Chile and Argentina) has also been hyperlinked within the application. Each of these narrations begins to play when this atlas subsection is initially loaded into a web browser, with the sound level set so the sound is inaudible. Positioning the cursor over either the shown hyperlinks defined by the text for one of these nations' names or that nation's claimed map sector will cause the sound level of the associated narration to be adjusted to maximum loudness. The text labels, since the cursor can only be over one at a time, can be used to render each narration aubible in isolation from the others and allows a user to listen to that perspective alone. By contrast, moving the cursor over one of the contested map sectors, because they partially overlap, will cause one, two, or all three of the narrations to be simultaneously audible, creating the aural sensation of competition among multiple perspectives and reinforcing the status of these sectors as the subject of dispute.

Although the use of these narrations in this map is informative, the primary motivation for the inclusion of the sounds in the design was to communicate the affective, emotional impact of the territorial contestation concerning Antarctica. A statement describing the motivation for the inclusion of these narrations as part of this map is included in the discussion of the map in the text panel to the right of the map.[11]

[11]This is shown in Figure B.5 but may not be reproduced very legibly so part of it is repeated

The *Territorial Claims* section of the *Cybercartographic Atlas of Antarctica* is available on the World Wide Web[12] and an audiovisual recording showing the atlas subsection concerning contested Antarctic territorial claims is included in the media examples on the CD-ROM accompanying this thesis (see Appendix G).

In preparation for the user evaluation sessions, the atlas section was modified such that the user was immediately presented with this subsection incorporating sound. This simplified the evaluation by avoiding having to guide the student through other atlas sections to find the part that used sound. This reduced the information presented to the evaluating user as part of their assesment of this atlas section but also focused them solely on one section of the atlas discussing territorial claims almost entirely in terms of contestation.

In addition, the version of the atlas used for the evaluation sessions was modified such that clicking on a national sector to display more information about a country's claim to territory would continue to display the map designed with embedded sounds. In the online version of the atlas, the maps associated with these country claim descriptions do not incorporate the sound design. In the context of evaluation sessions focused on sound in audiovisual maps, it was decided that presenting a user with visually identical maps, some with and some without sound, was an unnecessary complication. The video capture in Appendix G shows the version of the map used in the evaluation sessions.

B.3.2 Antarctic Exploration (AE)

The *Exploration* section of the atlas is structured around a narrative describing a mythical great southern continent called "The Island of Utopia" and the gradual replacement of this myth by increasingly detailed information collected during successive explorations of the region and Antarctica.[13] This story of explorations in Antarctica is represented, as shown in Figure B.6, using a combined display comprised of: 1) an horizontal array of interactive selectors, each associated with

here. "These disputes are conveyed through this map using audio files of vocal narrations triggered by moving your mouse over contested areas on the map. The function of vocal narration is to highlight the tension between these three countries." Although the quoted statement emphasizes the affective impact of the acoustic reinforcement of the contested claims, it is worth noting that the visual representation of multiple overlapped claims is also less common and consequently more likely to attract notice or comment than is the demarcation of disjoint territories. Complexity as a result of multiple claims to the same territory is not entirely absent from the visual representation in this atlas section.

[12]The *Territorial Claims* section of the atlas is available at http://atlases.gcrc.carleton.ca/antarctic/territorial/territories.xml.html. The subsection that highlights the contestation of territory discussed here is accessed by going to that web page and following the "sovereignty related conflicts" hyperlink at the bottom of the text panel.

[13]The chronology of Antarctic expeditions was based on McGonigal and Woodworth (2001, 380–503).

a leader of one or more expeditions to Antarctica and used to select an exploration journey for more detailed display;[14] 2) a small scale map of Antarctica that is updated according to the currently selected expedition; and 3) a scrollable multimedia panel used to provide additional information concerning each journey.

A user can cause the map to display the journey route, or routes, associated with an expedition leader by moving the cursor over the selector associated with that leader. This interactive control, if used to select each expedition leader in order, moving left to right, displays a chronology of the selected expeditions and approximates an animated version of the map produced using different software than that used for the final version of the atlas and included in the user evaluations.

By also clicking the mouse button while the cursor is positioned over one of the selectors, three things happen simultaneously. First, the multimedia panel is updated with information concerning the social, political, and scientific circumstances of the expedition. Second, the map display alternates between a view of those routes superimposed on a small scale relief map of Antarctica and a view of those same routes superimposed on an image intended to represent a point in time during a process of transformation starting with an artistic rendering of the mythical island and ending with the relief map. Whenever the cursor moves from being positioned over one expedition selector to another, the map display reverts to the transformation process imagery showing the routes of the newly selected expedition territorially diminishing the presence of the myth. Thus, sweeping the cursor across the array of selectors from earliest to latest visually replays the transformation beginning with the myth of the Island of Utopia and ending with a contemporary relief map of Antarctica. Third, a sound effect loop representing a mode of exploration essential to the expedition, becomes audible. The sound effects represent one of four possible modes of exploration: sea voyage, overland trek, airplane overflight, and satellite remote sensing.

The map display, multimedia panel, and sound effects are designed to reinforce each other. As Caquard et al. (2008) argued, pictograms in the route displays (ships under sail, footprints, an airplane, a satellite) were designed to assist users in deciphering the sound effects. Similarly, information presented in the multimedia panel in the form of explanation, images, videos, and narrations provide users with means to check or expand on information provided by the sounds or map images. The sounds, designed to suggest the "remoteness and isolation associated to the terrestrial exploration" (ibid., 1229) of Antarctica, also extend the information presented visually by, for example, rendering tactile the vibration of the wind that may have been experienced by an explorer trekking across the continent.

The script followed during evaluation sessions of this atlas section had the in-

[14]The interactive selector was often referred to within the atlas projects as a 'timeline'. Although the selector arranges the expeditions chronologically with later ones to the right of earlier journeys, the spacing of expeditions on the selector is uniform rather than proportional to the passage of time between expeditions and to verify that these are arranged in date order a user must examine each expedition's date, shown on the map display.

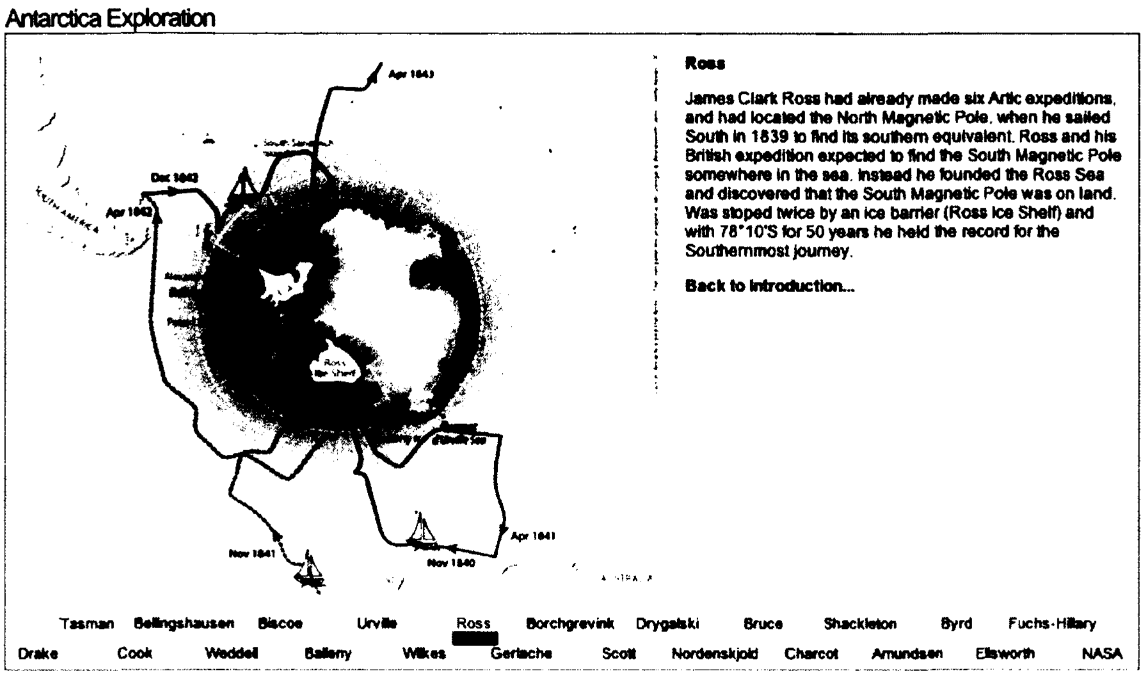


Figure B.6: Cybercartographic Atlas of Antarctica: Exploration (image courtesy of the Geomatics and Cartographic Research Centre, Carleton University).

terviewer ask each participant to begin by listening to a narration that described the myth of the Island of Utopia and then to sweep the cursor across the array of expedition selectors (see Appendix D). This, as discussed above, simulated the animated version of the map by having the user redisplay the expedition chronology and ensured that participants would discover the sounds used with the map.

The *Exploration* section of the *Cybercartographic Atlas of Antarctica* is available on the World Wide Web[15] and an audiovisual recording showing this section being used is included in the media examples on the CD-ROM accompanying this thesis (see Appendix G).

[15]The *Exploration* section of the atlas is available at http://atlases.gcrc.carleton.ca/antarctic/exploration/exploration.xml.html.

Appendix C: Ethics Application for User Evaluations

This section contains the main overview of the application made to the Carleton University Psychology Department's ethics review committee in preparation for conducting the user evaluations discussed in this thesis. Although this application refers to Appendices containing recruitment notices, consent forms, debriefing statements, and materials to be used during the evaluations, these have been removed here. Appendix D of this thesis contains the final versions of the forms and scripts used during the evaluation sessions and thus are more recent versions of those materials. Chapter 6 contains more detailed descriptions of the audiovisual maps that were the subject of thes evaluations than were included in the ethics application.

Usability Tests for Multimedia Maps that Use Sound

Description of Purpose

The proposed research seeks to examine users perceptions of the usability and usefulness of multimedia cartography examples that combine sound with visual elements (graphics, images, and text) and to examine users' interpretations and understandings of information presented using these examples. The evaluation is exploratory. The reactions of participants to the examples, the ease or difficulty they experience in working with and understanding the materials are very important to us as are their abilities to demonstrate their understandings of the presented information. The main questions to be answered through the evaluations are:

1. Are users able to work with the materials? Do users understand how the sounds relate to other media elements included in the map examples? Do they require assistance in working with the examples and, if so, in what ways?

2. Based on answers to question 1, do users understand what is being presented? Are they able to express information gained through working with the mapping examples or answer questions based on those materials?

3. How do users react to the mapping examples? How do they work with the examples and what do they say about the examples after having worked with them?

Description of Methods

For this study, we will present four multimedia cartography examples to the participants and ask them to work with each example, discuss or demonstrate their understanding of the information presented, and ask them to evaluate aspects of the design of each example. There are no potential risks or discomforts in this study.

Participants

Approximately 10-15 participants will be recruited for the study and each participant will receive course credit. Data collection will be done in a manner that ensures participant anonymity. Data collection forms will be coded and participant names will not be included on the data collection forms. Only researchers working on the study will have access to the data.

PSYC 1001 and 1002 students shall be recruited to participate in the study and course credit will be given (.5% per half hour and it is anticipated that an evaluation session with one participant should last about an hour). Recruitment will be conducted using the SONA system. The recruitment information to be used to advertise the study in SONA is included as Appendix A of this application. Additionally, a small number of acquaintances of the researchers shall be recruited by giving them a written description of the study for recruitment purposes, either delivered in person or through email, and they will be paid $10 for their participation. Appendix B contains the recruitment notice that will be distributed in person or through email. As outlined above, coding of the collected data will be used to protect the anonymity of the participants, regardless of how they are recruited.

Participants will be asked, at the beginning of the evaluation session, to sign an informed consent statement and a form consenting to the collection of video and audio recordings of their evaluation session. The informed consent form is included as Appendix C of this application and the video and audio consent form is included as Appendix D.

Each participant will be presented with a debriefing statement at the end of the evaluation session. This statement will provide them with information concerning the purpose of the study and who may be interested in the results. The debriefing statement will also provide references for further information and a set of contact names for the participant if they have questions about the research or ethical concerns about the research. The debriefing statement is included as Appendix E of this application.

Materials

The materials to be used during the evaluation sessions include four multimedia cartography applications that participants will be asked to work with. This application contains the following material describing these applications:

1. Appendix F of this application describes each of these examples in more detail.

2. This application includes a CD-ROM containing:

    i. Video captures of the Antarctic Exploration, Antarctic Territorial Claims, and Election maps: these are all of the maps that include vocal narrations. These videos can be viewed using QuickTime and show how sounds are used within the maps.

    ii. The audio files for the narrations (in Windows WAV format). These include the narrations used in the scenarios for the Antarctic Exploration, Antarctic Territorial Claims, and Election maps. These audio files contain the complete narrations used in the maps.

    iii. Written transcripts for the vocal narrations used in the Antarctic Territorial Claims map (in Microsoft Word format).

    iv. The synthesized music (MIDI) composition that is used as part of the Trade Data Browser. This file can be played using QuickTime.

The researchers conducting the evaluations will conduct the evaluations using prepared interview scripts and questionnaires. These scripts are included in Appendix G.

Procedure

1. As outlined above, PSYC 1001 and 1002 students will be recruited to participate in this study using SONA.

2. Participants will be introduced to the study and told they will receive credit according to the established policy of 0.5% per half hour experiment participation.

3. Participants will then be given an informed consent form to sign (See Appendices B).

4. Each participant will then commence with the tasks in the usability test study on the computer. After some brief introduction to the evaluation tasks, the participant will be asked to consent to having the researchers audio record the discussions and video record the computer screen during the evaluation (the Audio and Video Recording Consent Form is attached as Appendix D). They will be presented with four examples of multimedia maps that use sound (see Appendix F for more information concerning each of the examples) and will be given specific instructions for each example. They will be asked to work with the example and then to either explain their understanding of some aspects of the information presented in the example or to answer specific questions by working with the example. Tasks and scenarios are provided in Appendix G. Participants will be allowed to take rest breaks throughout the study.

5. Participants will then be debriefed as shown in Appendix E.

Appendix D: User Evaluation Scripts and Forms

This section contains consent forms, debriefing statements, and the script followed by interviewers during user evaluation sessions concerning audiovisual maps discussed in this thesis. Note that to complete each participant's session within an hour only three maps, selected beforehand by the interviewers, were considered by each participant. For each paticipant, the section of these scripts specific to the other map would be omitted.

Informed Consent Form: Usability Tests for Multimedia Maps that Use Sound

Introduction:

The purpose of an informed consent is to ensure that you understand the purpose and your involvement in the study. The informed consent must provide enough information so that you can determine whether or not you wish to participate in this study.

Study: Usability Tests for Multimedia Maps that Use Sound

Research Personnel:

The following personnel are involved in this research project and may be contacted at any time.

- Principal Investigators:
    - Glenn Brauen [redacted]
    - Shelley Roberts [redacted]
    - Aren Hunter [redacted]
- Faculty Sponsor: Dr. Avi Parush, 613-520-2600 ext. 6026, avi_parush@carleton.ca.

If any ethical concerns about this study should arise, please contact Dr. Shelley Parlow (Ethics Committee for Psychological Research), 613-520-2600, ext. 2691, shelley_parlow@carleton.ca. Should you have any other concerns about this study, please contact Dr. Anne Bowker (Chair, Dept of Psychology,) 613-520-2600, ext. 2648, psychchair@carleton.ca.

Purpose:

The purpose of this study is to examine users perceptions of the usability and usefulness of multimedia cartography examples that combine sound with visual elements (graphics, images, and text) and to examine users' interpretations and understandings of information presented using these examples.

Task Requirements:

You will be asked to complete tasks with multimedia map examples. You will then be asked to answer questions based on information presented in the examples or explain your understanding of the examples

Duration and Locale:

The duration of this study will be one hour. Participants will earn course 1% course credit for their participation.

Potential Risk or Discomfort:

There are no potential risks or discomforts in this study. Participants will be allowed to take rest breaks if they require by asking the researchers at any time during the study.

Confidentiality:

All participants' data will be coded and remain anonymous. The data collected will be coded such that your name will not be associated with the data. The data is to be used only by the researchers involved in this project.

Right to Withdraw:

You have the right to withdraw from the study at any time without penalty.

I have read the above description of the study and understand the conditions of my participation. I agree to participate in this research project.

Participant Name: ______________________________

Participant Signature: ______________________________

Witness/Researcher Name: ______________________________

Witness/Researcher Signature: ______________________________

Usability Tests for Multimedia Maps that Use Sound
Video and Audio Recording Consent Form

We would like to record information about how you work with the multimedia map examples for the purposes of evaluating the software and showing the results of our testing. This audio and video taping consent form requests your permission for us to audiotape your statements and videotape the monitor while you are performing the tasks.

By signing this form you give your consent to use the information you give us, your voice and verbal statements, and video capture of your computer screen, but not your face and not your name, only for the purpose of evaluating the software and showing the results of our testing.

Name: ______________________________
Signature: ______________________________
Date: ______________________________

Demographics Questionnaire

1. Gender: M F

2. Age:_______

3. Please rate your fluency in English (1=poor, 5=fluent)

i.	Reading:	1	2	3	4	5
ii.	Writing:	1	2	3	4	5
iii.	Speaking:	1	2	3	4	5

4. Do you consider yourself fluent in any other languages and if so which ones?

____

____

5. Please state your current education level and program: (e.g. 1st year Psychology)

____

6. Do you have any hearing impairments? If yes, please explain.

____

7. Is your vision normal or corrected to normal (i.e. you have glasses and are wearing them=corrected to normal?) Yes No

8. Are you colour blind? If, yes please explain_________________________

9. Have you ever studied a musical instrument or voice? Yes_____ No______

   1. If yes, please give a brief description of your studies (e.g. 4 years of piano from age of 5 to 9 with one-on-one instruction and passed grade 2 conservatory.)

____

10. Please rate the amount of time you listen to music (1=never, 5=very often)

    1 2 3 4 5

11. Briefly explain your audio listening habits: (e.g. listen to my mp3 player on the way to school, listen to audio books or podcasts, only listen to music).

____

____

12. Have you ever voted in a provincial or federal election? Yes_____ No_______

13. Do you know what your home ridings are? Yes_____ No_______

a) If yes please name your;

Federal electoral district_________________________________

Provincial electoral district______________________________

14. Do you have an interest in Canada's economic issues? Yes_____ No_______

15. Do you have an interest in Canada's trade with the world? Yes_____ No______

a. if yes, please rate your level of knowledge for Canada's trade issues (1=not much knowledge at all, 5=extensive knowledge)

1 2 3 4 5

Usability Test for Multimedia Maps that Use Sound

Facilitator's Notes:

- Have the mapping examples ready and seat the participant at the computer.
- Read the following to the user:

Introduction:

Thank you for agreeing to participate in this evaluation of multimedia map examples that use sound. We look forward to hearing your comments and seeing your reaction to the mapping examples we are going to show you. You will work with several example maps and asked to describe your understanding of the example and work with the example to answer specific questions. Please speak out your thoughts as you work with the examples as much as possible. For example, state what you are trying to find as you work with an example and whether or not you find things where you expect them. If you find a map difficult to understand or use, say so. Remember that we are testing the map examples – not you – so don't be afraid to state your opinion about the examples. Feel free to ask questions at any time.

As a warm-up and to get you used to speaking out loud as you work with the examples, I'm going to ask you to work with a web-page briefly:

- I've loaded www.carleton.ca in the web browser for you.
- Please find the registrar's webpage and look for information about student exchanges. Remember, the purpose of this exercise is for you to speak your thoughts concerning what you are looking for, whether or not you can find it, whether or not you find something else relevant instead, etc. Do not worry about whether or not you have difficulty with the assigned task. We want to understand if there is something about the task or website that makes it difficult for you.

(*Allow the participant to briefly search for the requested information and encourage them to voice their thoughts. But this is just a warm-up so keep it as brief as possible.*)

I am going to give you an idea of the types of things I am looking for:

- If you don't understand how to work with a map example, please say so.
- If you don't understand the map (visuals or the sounds in the map), please say so or ask questions to clarify your understanding.
- If you read a description in one of the examples but don't understand it, please say so.

ELECTION MAP

(Display the election map web page on the computer but do not cause any sounds to start. Ask the participant to examine the page and observe their initial work with the example and their reactions. If they have difficulty with the example, answer their questions but note the difficulties.)

Instructions: Please take a few minutes and play with or explore the website. Remember to think aloud as you explore.

1. Usability

Time taken to explore: __________

While users explore the map take notes on the following:

- Are users able to work with the materials?
- Do users understand how the sounds relate to other media elements included in the map examples?
- Do they require assistance in working with the examples and, if so, in what ways?

2. Usefulness

1. Describe the sounds used in the map. How many sounds are there and how do they relate to the visual map?
   (Purpose of the task: to understand whether or not the participant understands the relationship between the sounds and the visual map (electoral districts and legend elements).

   ____

   ____

   ____

   ____

2. Compare the election results for Argenteuil-Mirabel and Ottawa-Centre (62).
   (*Note whether or not the participant discusses the apparent difference in how close the results in the two ridings were. If not, explicitly ask a follow-up question.*)

   ____

   ____

   ____

   ____

   a. Which district do you think had the closer election results?

      ____

   b. How certain are you of your answer on a scale of 1 to 5? (1=extremely certain, 5=not certain at all)

      1 2 3 4 5

3. Do you think the Liberals consider the election to have been a success?

   Yes No

   a) What makes you think so?

      ____

      ____

(*The facilitator may have to point the participant to the legend to allow them to isolate individual speeches but let them try on their own first.*) *(Purpose of the task: to understand whether or not they pay attention to the content of the leaders' speeches or just the mixing of the volume levels.)*

Post Questions:

1. Did you notice the different volume levels of individual leader's speeches as the mouse moves over the districts?
2. Did you relate the volume levels to that of the election results?
3. Do you understand how ridings/parties/federal election work?
4. Do you know which federal district Carleton University would fall under?

3. Satisfaction Questionnaire: Election Map

Instructions: Please answer the following usability questions based on the map that you just interacted with. Your answers are very important as they will aid in the improvement of these maps for future use.

1. My overall experience with the map can best be described as:

Terrible Indifferent Wonderful
1 2 3 4 5

2. When I had to extract information from the map to answer questions I generally felt:

Frustrated Indifferent Satisfied
1 2 3 4 5

3. I general I felt the map was:

Dull OK Stimulating
1 2 3 4 5

4. Answering questions based on the map was:

Difficult Moderate Easy
1 2 3 4 5

5. I would describe the amount of visual information on the map as:

Inadequate Adequate Too much
1 2 3 4 5

6. I would describe the amount of auditory information on the map as:

Inadequate Adequate Too much
1 2 3 4 5

7. Learning to use the system was:

Difficult OK Easy
1 2 3 4 5

8. I used the **auditory** information in the map to answer the question:

Never Sometimes Always
1 2 3 4 5

9. I felt the **auditory** information in the map helped me answer the questions:

Never Sometimes Always
1 2 3 4 5

10. I felt that the **auditory** information made using the map more interesting:

Disagree		Somewhat		Agree
1	2	3	4	5

11. I felt that the **auditory** information was not helpful but very distracting:

Disagree		Somewhat		Agree
1	2	3	4	5

12. Would you want to use this map again:

No		Maybe		Yes
1	2	3	4	5

13. What did you like most about the map?

____

____

14. What did you least like about the map?

____

____

15. Do you have any suggestions for improving this map?

____

____

Antarctic Territorial Claims Map

(Facilitator: Display the territorial claims section of the Atlas of Antarctica on the computer but do not cause any sounds to start. Ask the participant to examine the page and observe their initial work with the example and their reactions. If they have difficulty with the example, answer their questions but note the difficulties.)

Instructions: Please take a few minutes and play with or explore the website. Remember to think aloud as you explore.

1. Usability

Time taken to explore: __________

While users explore the map please take notes on the following:

- Are users able to work with the materials?
- Do users understand how the sounds relate to other media elements included in the map examples?
- Do they require assistance in working with the examples and, if so, in what ways?

2. Usefulness

1. Describe the sounds used in the map. How many sounds are there and how do they relate to the visual map?

   ____

   ____

   ____

   ____

   (Purpose of the task: to understand whether or not the participant understands the relationship between the sounds and a) the territorial claims on the visual map, and b) the hyperlinked country names in the text panel.)

2. Why do only some of the territories have associated sounds?

   ____

   ____

   ____

   ____

   (Purpose of the task: to further understand whether or not the participant understands the role of sound in highlighting only the claims for contested territory and that these sounds therefore only apply to some of the territories on the map and to some of the countries with Antarctic claims.)

3. Compare the map section for New Zealand to that of Chilé.
   a) What are the main differences you see?

      ____

      ____

   b) What do you think they mean?

      ____

      ____

   (Purpose of the task: to encourage the participant to work with the map so that they discover the sounds built into the map and to see how they react to presence of no sounds, one sound, or multiple sounds depending on where they position the cursor)

4. Do the sounds in this section make sense to you? Yes No
   a) Describe how the sound is related to the graphic elements (text, map graphics) of the atlas in this section.

      ____

      ____

3. Satisfaction Questionnaire: Antarctic Territorial Claims Map

Instructions: Please answer the following usability questions based on the map that you have just interacted with. Your answers are very important as they will aid in the improvement of these maps for future use.

1. My overall experience with the map can best be described as:

Terrible		Indifferent		Wonderful
1	2	3	4	5

2. When I had to extract information from the map to answer questions I generally felt:

Frustrated		Indifferent		Satisfied
1	2	3	4	5

3. I general I felt the map was:

Dull		OK		Stimulating
1	2	3	4	5

4. Answering questions based on the map was:

Difficult		Moderate		Easy
1	2	3	4	5

5. I would describe the amount of visual information on the map as:

Inadequate		Adequate		Too much
1	2	3	4	5

6. I would describe the amount of auditory information on the map as:

Inadequate		Adequate		Too much
1	2	3	4	5

7. Learning to use the system was:

Difficult		OK		Easy
1	2	3	4	5

8. I used the **auditory** information in the map to answer the question:

Never		Sometimes		Always
1	2	3	4	5

9. I felt the **auditory** information in the map helped me answer the questions:

Never		Sometimes		Always
1	2	3	4	5

10. I felt that the **auditory** information made using the map more interesting:

Disagree		Somewhat		Agree
1	2	3	4	5

11. I felt that the **auditory** information was not helpful but very distracting:

Disagree		Somewhat		Agree
1	2	3	4	5

12. Would you want to use this map again:

No		Maybe		Yes
1	2	3	4	5

13. What did you like most about the map?

____

____

14. What did you least like about the map?

____

____

15. Do you have any suggestions for improving this map?

____

____

Antarctic Exploration Map With Sound

(*Facilitator: Display the exploration section of the Atlas of Antarctica on the computer but do not move the cursor over the timeline to avoid playing the timeline sounds. Do not allow the participant to freely examine the example, instead* ***play the narration explaining the myth of the "Island of Utopia". Then immediately ask them to move on to Task 1***.)

1. **Usability**

Time taken to explore: __________

While users interact with the map please take notes on the following:

- Are users able to work with the materials?
- Do users understand how the sounds relate to other media elements included in the map examples?
- Do they require assistance in working with the examples and, if so, in what ways?

2. Usefulness

a) Please sweep the mouse slowly across the timeline without clicking the buttons. Explain what you think the sounds indicate?

____

____

____

(*If they have difficulty with this task or are unsure, ask them to take more time to examine the timeline – this time clicking on elements of the timeline, if they wish – and then ask them again to explain the sounds.*)

b) What has been the predominant mode of transportation used to explore Antarctica?

____

(*Note if the participant works with the map to answer the question and, if so, how.*)

c) Can you tell me who first explored Antarctica on foot?

____

(again inquire about how they obtained that information sound, text, auditory)

d) Did you notice that the routes in the visual displays changed for each exploration?

Yes No

3. Satisfaction Questionnaire: Antarctic Exploration Map With Sound

Instructions: Please answer the following usability questions based on the map that you just interacted with. Your answers are very important as they will aid in the improvement of these maps for future use.

1. My overall experience with the map can best be described as:

Terrible		Indifferent		Wonderful
1	2	3	4	5

2. When I had to extract information from the map to answer questions I generally felt:

Frustrated		Indifferent		Satisfied
1	2	3	4	5

3. I general I felt the map was:

Dull		OK		Stimulating
1	2	3	4	5

4. Answering questions based on the map was:

Difficult		Moderate		Easy
1	2	3	4	5

5. I would describe the amount of visual information on the map as:

Inadequate		Adequate		Too much
1	2	3	4	5

6. I would describe the amount of auditory information on the map as:

Inadequate		Adequate		Too much
1	2	3	4	5

7. Learning to use the system was:

Difficult		OK		Easy
1	2	3	4	5

8. I used the **auditory** information in the map to answer the question:

Never		Sometimes		Always
1	2	3	4	5

9. I felt the **auditory** information in the map helped me answer the questions:

Never		Sometimes		Always
1	2	3	4	5

10. I felt that the **auditory** information made using the map more interesting:

Disagree		Somewhat		Agree
1	2	3	4	5

11. I felt that the **auditory** information was not helpful but very distracting:

Disagree		Somewhat		Agree
1	2	3	4	5

12. Would you want to use this map again:

No		Maybe		Yes
1	2	3	4	5

13. What did you like most about the map?

____

____

14. What did you least like about the map?

____

____

15. Do you have any suggestions for improving this map?

____

____

Canada's Trade with the World

(Facilitator: Before showing the participant the trade data browser, present the participant with a copy of the one-page description of trade terminology and the trade variables presented in the trade browser. Let them read it and answer any questions they have about it. Then display the trade data browser on the computer with the "learning mode" displayed). Ask the participant to examine the page and observe their initial work with the example and their reactions. If they have difficulty with the example, answer their questions but note the difficulties.)

1. **Usability**

   Time taken to explore: __________

   While users interact with the map please take notes on the following:

- Are users able to work with the materials?
- Do users understand how the sounds relate to other media elements included in the map examples?
- Do they require assistance in working with the examples and, if so, in what ways?

2. Usefulness

(in "learning mode")

1. Do you understand how the individual sounds are related to the audio indicators?

   ____

   ____

2. How many audio indicators are there?

   ____

   ____

(Allow the participant a reasonable amount of time to work with this but, if they are stuck, help them to understand the relationship between the sounds and the visuals so that they can continue with the subsequent tasks).

(still in "learning mode")

3. Compare the regions "Mexico, Central America, and Caribbean" to "South America". Which region do you think Canada predominantly exports to and which region do you think Canada predominantly imports from?

   ____

   ____

(switch the trade browser to "default use mode" now)

4. Compare the former U.S.S.R. and Europe. Which region do you think Canada does more trade with?

   ____

   ____

5. What makes you think so and how sure are you?

   ____

   ____

6. How many regions does Canada have balanced trade with? How sure are you?

   ____

   ____

3. Satisfaction Questionnaire: Canada's Trade with the World

Instructions: Please answer the following usability questions based on the map that you just interacted with. Your answers are very important as they will aid in the improvement of these maps for future use.

1. My overall experience with the map can best be described as:

   Terrible Indifferent Wonderful
   1 2 3 4 5

2. When I had to extract information from the map to answer questions I generally felt:

   Frustrated Indifferent Satisfied
   1 2 3 4 5

3. I general I felt the map was:

   Dull OK Stimulating
   1 2 3 4 5

4. Answering questions based on the map was:

   Difficult Moderate Easy
   1 2 3 4 5

5. I would describe the amount of visual information on the map as:

   Inadequate Adequate Too much
   1 2 3 4 5

6. I would describe the amount of auditory information on the map as:

   Inadequate Adequate Too much
   1 2 3 4 5

7. Learning to use the system was:

   Difficult OK Easy
   1 2 3 4 5

8. I used the **auditory** information in the map to answer the question:

   Never Sometimes Always
   1 2 3 4 5

9. I felt the **auditory** information in the map helped me answer the questions:

   Never Sometimes Always
   1 2 3 4 5

10. I felt that the **auditory** information made using the map more interesting:

Disagree		Somewhat		Agree
1	2	3	4	5

11. I felt that the **auditory** information was not helpful but very distracting:

Disagree		Somewhat		Agree
1	2	3	4	5

12. Would you want to use this map again:

No		Maybe		Yes
1	2	3	4	5

13. What did you like most about the map?

____

____

14. What did you least like about the map?

____

____

15. Do you have any suggestions for improving this map?

____

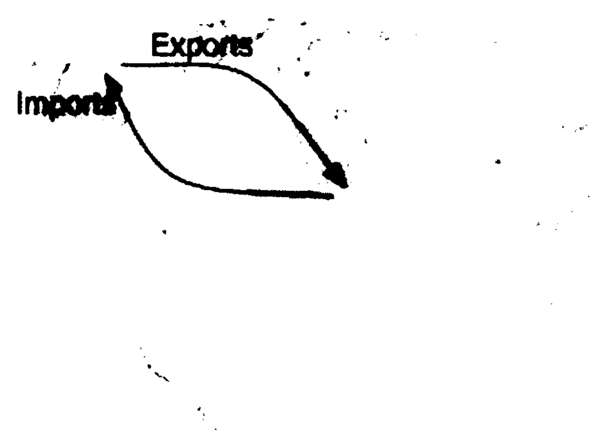
____

Information Presented in the Trade Data Browser

The trade data browser presents information concerning the voluntary exchange of goods and/or services – often referred to as "***trade***" – between Canadians, Canadian businesses, governments, and institutions with people, businesses, and organizations in other regions of the world. All of the data presented relates to the trade of goods and services in to or out of Canada as shown in the figure below.

The inter-regional trade totals presented in this browser are made up of very many individual transactions in which goods or services are sold from Canada to other regions or are sold from other regions into Canada. For this data browser, trade transactions are defined as follows:

- ***Exports***: When Canadians sell goods and/or services to people or organizations in other regions, these goods and services are ***exported*** from Canada and are called ***exports***.
- ***Imports***: when Canadians buy goods and/or services from people or organizations in other regions, these goods and services are ***imported*** into Canada and are called ***imports***.



Usability Tests for Multimedia Maps that Use Sound
Debriefing Statement

The research we are conducting is of interest to researchers designing multimedia and multisensory maps and atlases. Carleton University's *Geomatics and Cartographic Research* Centre (GCRC) is one centre for this work but interest in attempting to understand how well people can understand and work with maps that engage multiple sensory modes (e.g., hearing and touch as well as sight) exists internationally. The GCRC, under the direction of Dr. D. R. Fraser Taylor, has been working to develop theoretical and practical guidelines for the development of these types of mapping products. Some of the examples you worked with during this evaluation are from atlases developed by the GCRC and partner organizations, including Carleton University's Human Oriented Technology Lab (HOTLab). For more information:

- Taylor, D. R. F. (ed.). 2005. *Cybercartography: Theory and Practice*, Amsterdam: Elsevier.
- Taylor, D. R. F., and S. Caquard (Guest eds.). 2006. Special Issue on Cybercartography, *Cartographica*, 41:1.

The present study was designed to examine users' perceptions of the usability and usefulness of multimedia cartography examples that combine sound with visual elements (graphics, images, and text) and to examine users' interpretations and understandings of information presented using these examples. Your feedback will be very important in helping us to understand whether or not these atlas designs are effective and which of the design ideas are more or less successful in helping atlas users to understand complex issues. Thank you for participating in the study. Your time and effort are greatly appreciated!

Please feel free to contact us at any time if you wish to discuss any aspects of the research.

- Principal Investigators:
  - Glenn Brauen [redacted]
  - Shelley Roberts [redacted]
  - Aren Hunter [redacted]
- Faculty Sponsor: Dr. Avi Parush, 613-520-2600 x6026, avi_parush@carleton.ca.

If any ethical concerns about this study should arise, please contact Dr. Shelley Parlow (Ethics Committee for Psychological Research), 613-520-2600 x2691, shelley_parlow@carleton.ca. Should you have any other concerns about this study, please contact Dr. Anne Bowker (Chair, Dept of Psychology,) 613-520-2600 x2648, psychchair@carleton.ca.

Appendix E: User Evaluation Analysis Details

This Appendix presents details concerning the four audiovisual map examples evaluated during sessions designed to generate data concerning the *usability* and *usefulness* of the examples and, specifically, the usefulness of sound as part of the audiovisual designs, as discussed in Chapter 6, along with details of the analysis of that data. Section E.1 presents a summary of the cross-sectional analysis codes assigned, as discussed in Chapter 6. Section E.2 presents detailed analysis of the sessions based on that coding summary.

Throughout the discussion, as in Chapter 6 and Table 6.2, the main analysis discussion points are designated and referenced as 'ad.n' and specific participants are designated as 'Pn' where n shows a participant's position in the chronological ordering of the evaluation sessions.

E.1 Map Evaluation Data: Cross-sectional Codes

Table E.1 provides a summary of cross-sectional classification codes assigned to text segments of user evaluation sessions transcripts during analysis. A check mark in the table indicates that the corresponding code (table row) was assigned at least once to a segment of text in the transcript of that user's evaluation of the listed audiovisual map example (table column). This coding summary was developed and used as a guide while writing the detailed analysis discussions presented in Section E.2.

Table E.1: Assignment of cross-sectional codes to participant evaluations of map examples. A marked session indicates that the code was assigned to at least one selection of text in the transcript for the participant's evaluation of that map example. Column codes indicate the map being evaluated and are ordered, left-to-right, in order evaluated by each participant: E (election), T (trade), C (territorial claims), X (exploration). Continuing on next page...

		Participants and Ordered Map Example Codes																							
		P1			**P2**			**P3**			**P4**			**P5**			**P6**			**P7**			**P8**		
Categories	**Codes**	X	T	E	T	C	X	C	E	T	E	X	C	X	C	E	T	E	C	C	X	T	E	T	X
Assistance	given				✓					✓							✓	✓				✓		✓	
Interpretation	both sound and visuals used to interpret	✓					✓	✓			✓	✓		✓					✓	✓	✓				✓
	based on sound effects						✓					✓		✓		✓					✓				
	using visual information INSTEAD of audio				✓		✓		✓	✓	✓	✓	✓		✓	✓	✓	✓				✓			✓
	sound unacknowledged as informing answer											✓													✓
	sounds explains visuals	✓						✓				✓													
	visuals explain sounds	✓	✓	✓		✓		✓	✓			✓	✓				✓	✓	✓		✓	✓			
	lack of confidence expressed							✓			✓						✓	✓							
	asked: purpose of sound			✓	✓								✓	✓											
Enjoyment	approval expressed	✓	✓	✓					✓					✓	✓					✓				✓	

Table E.1: Assignment of cross-sectional codes to participant evaluations of map examples.
Continued from previous page...

		Participants and Ordered Map Example Codes																							
		P1			P2			P3			P4			P5			P6			P7			P8		
Categories	**Codes**	X	T	E	T	C	X	C	E	T	E	X	C	X	C	E	T	E	C	C	X	T	E	T	X
	disapproval expressed	✓		✓		✓					✓		✓		✓	✓						✓			
	repetitive	✓		✓																					
	surprise		✓		✓			✓							✓								✓	✓	
	notes aesthetic contribution of sound													✓											
Listening	didn't initially notice acoustic changes				✓																	✓			
	lack of confidence (audio discrimination)				✓					✓															
	prompt to listen required	✓																							
	qualitative sound changes detected						✓					✓		✓							✓				✓
	unanticipated sound distinctions	✓										✓	✓												✓
Use of speech	attending to speech content			✓		✓		✓	✓		✓		✓		✓	✓		✓		✓			✓		
	uncomfortable w/ overlapped speech			✓		✓		✓			✓		✓		✓	✓							✓		

Table E.1: Assignment of cross-sectional codes to participant evaluations of map examples.
Continued from previous page. . .

		Participants and Ordered Map Example Codes																							
		P1			P2			P3			P4			P5			P6			P7			P8		
Categories	**Codes**	X	T	E	T	C	X	C	E	T	E	X	C	X	C	E	T	E	C	C	X	T	E	T	X
	overlapped speech: symbolic interpretation					✓			✓				✓					✓	✓	✓					
	overlapped speech: no interpretation														✓										
	detected: quantitative use of speech								✓		✓					✓		✓							
	undetected: quantitative use of speech			✓																			✓		
	Utopia narration played then ignored	✓					✓					✓		✓							✓				✓
	two voices means translation					✓																			
Trade variables and dynamic legends	detected: data-based mixing of instruments		✓		✓																			✓	
	undetected: quantitative use of music									✓							✓					✓			
	successfully used trade indicators w/o legends		✓		✓					✓												✓		✓	

Table E.1: Assignment of cross-sectional codes to participant evaluations of map examples. Continued from previous page...

		Participants and Ordered Map Example Codes																							
		P1			P2			P3			P4			P5			P6			P7			P8		
Categories	**Codes**	X	T	E	T	C	X	C	E	T	E	X	C	X	C	E	T	E	C	C	X	T	E	T	X
	used visual legends: one or more of final Qs									✓							✓								
	unable to answer: one or more of final Qs																✓							✓	
	unclear about legends as controls (used before map)				✓																				
	using legends as controls w/o prompting		✓		✓																	✓		✓	

E.2 Detailed User Evaluation Session Analysis

E.2.1 Ottawa Area Federal Election Sound Map: June 28, 2004 (OAFESM)

(ad.1) Four of six (4/6) participants decided that the mixing of speeches was significant: While working with *Ottawa Area Federal Election Sound Map: June 28, 2004* (OAFESM), most participants interpreted the overlapped speech design as significant, although alternative meanings were ascribed to the use of sound. The intended quantitative interpretation of the varied gain playback of each leader's speeches was understood by most participants, possibly through an interpretation of both visual and acoustic information provided in the application (ad.3 on p.324). Two participants, P3 and P6, in addition to the quantitative interpretation, said they thought the overlapped speeches represented "the [election] debates."[1] Finally, two participants did not indicate that they thought the overlapped speech design was either meaningful or necessarily intentional.

P1, while initially working with OAFESM, described the overlapping speech design as "overwhelming" and never indicated that she understood the quantitative representation. Subsequently, in response to interviewer questions, P1 stated that she had noticed the volume levels of the speeches varying as she moved the cursor over the map and that she *had not* related those changing sound levels to the election results.

Similarly, P8 noticed the varying sound levels of the leaders' speeches as she moved her cursor over the map and she was able to hear and determine that the speeches were composed of multiple overlapping sound clips. However, rather than interpreting that concurrent sound levels were designed to indicate a relative ranking of election results for the selected region, she seemed to interpret the loudest one as being that of the winning party and stated that she did not understand why she was hearing "other voices that are playing in the background." She seemed to interpret this as a defect in the map design and, when asked, also stated that she did not relate the varying speech sound levels to the election results.

(ad.2) All participants commented on sampled speech content and three of six (3/6) disliked that they did not know how to isolate each speech: Without exception, participants working with OAFESM paid attention to, or attempted to pay attention to, the spoken content of the sampled speeches. The transcripts of the user evaluation sessions show this in a number of ways: 1) expressions of disapproval by participants with not being able to isolate individual speeches; 2)

[1]This overlapped sound design was chosen in part because the simultaneously playing speeches of multiple politicians was reminiscent of broadcasts of Question Period from the House of Commons, in which the parliamentarian currently recognized by the Speaker of the House is rarely allowed to pose or answer a question without other members interrupting throughout.

concerns expressed about perceived inconsistencies between visual and acoustic information included in the application; and 3) participants "talking back" to the map, making comments in response to what they were hearing, or answering questions about election results based on what they heard in the speeches.

First, as discussed above (ad.1 on p.323), P1 described the overlapping speeches of OAFESM as "overwhelming." When discussing the application with the interviewer, P1 stated "I still don't really like how it has more than one speech going at the same time because [...] I have to hone in on one specific thought, and they are all kind of jumbled in together." Similarly, P4 commented on the use of overlapping recordings of speech with the map and contrasted that to the use of the map legend to isolate individual leader's speeches, saying "Yeah, I'm having a hard time understanding each individual one but when I go to the right-hand side [legend] which has each individual leader speaking, it's easier to pay attention that way obviously because there's only one thing to focus on." Upon first hearing the overlapped speeches, P5 asked "Why am I getting two sounds?"

Second, as discussed above (ad.1 on p.323), P1 detected the varying loudness of the speeches at least in aggregate but did not interpret this as quantitatively related to the election results. Without that understanding, she focused on what she decided was an inconsistency in the design of the map when, upon moving the cursor over a region coloured to indicate that the Liberals had won the election, she heard the recorded comment by Stephen Harper, the Conservative party leader, that "it's time to break the Liberal cycle [...]" She commented more than once that this made the map seem incorrect. Therefore, she thought the map design should not have included negative comments about the Liberals, obviously coming from someone other than the Liberals, while the cursor was positioned over a region in which they had won the election. Rather she was expecting a sound design that mirrored the visual distinctness of the map.

Third, most participants mixed running commentary on the recordings into their *think aloud* verbalizations concerning their work with OAFESM. Some commented on the use of mixed English and French in the recordings (P3 and P4). P5 stated that he thought one of the speakers had "a really annoying voice" and responded to some of the policy statements included in the speeches that "I disagree with these guys, but that's not the point." Despite the map legend not identifying individual leaders, participants also commented that they recognized the voices of leaders, with at least one participant correctly naming each of Paul Martin (P6) and Jack Layton (P5). P8, in response to an interviewer question concerning whether or not the Liberals would have considered the election a success, referred to Paul Martin's opening comment from the recordings, saying that "[t]hey had lost many votes."

(ad.3) Participants used both visual and acoustic information to interpret the sound design: The participants working with OAFESM used both visual and acoustic information in deciding how to interpret the sound design. Having already used the legend colour blocks to isolate speeches, P1 was asked how many separate speeches there were. She noted that because of the relatively short loops, you could hear repetition and become quickly familiar with what was said in each. She also referred to the number of colour blocks in the legend, stating that "you can definitely tell that it's different people speaking also because [...] it shows the parties," thereby applying the visual differentiation of the party colour scheme to reinforce the distinctiveness of each speech recording.

When P3 initially used OAFESM, she listened to the mixing of the party leaders' speeches as she moved the cursor over the map and commented on characteristics of the recordings (e.g., discussing with the interviewer whether or not the recordings were the sound of "the debates" and commenting on use of both French and English in the recordings). Before beginning to ask the scripted questions, the interviewer suggested that P3 take another minute and explicitly suggested that she could, if she wanted, read the instructions for the use of the map which she had not yet done. Subsequently, the cursor paused near the text briefly, the cursor then resumed moving between regions on the map, and then P3 asked "Is the one who talks the loudest, the one that wins?" Despite her initial comments expressing approval concerning the audiovisual design, there was no indication that P3 was interpreting the sound design as quantitatively related to the data until after she read the text.

During his initial examination of the application, P6 could be heard to discuss the continuity of "one voice" (Paul Martin, leader of the Liberals) as he moved the cursor between map regions, and to comment on the addition of Stephen Harper's speech as the cursor moved from Québec ridings, across the Ottawa River, and into Ontario electoral districts. P6 was thus describing his attention to the acoustic continuity within at least the Martin speech to determine that the auditory changes he heard, corresponding to visual region highlighting updates, constituted changes in *how* the audio loops were being played back rather than changes in *what* was being played back. In the midst of this dialogue, he said "Oh, I can see on the side of the map it indicates the different parties that won in these areas, so perhaps [those are the] parties that are speaking when I'm over those areas." Acoustic continuity within recordings then indicated to P6 that the speeches were applied across the entire map surface while the visual legend helped him to interpret what each speech represented.

(ad.4) Three of six (3/6) participants successfully compared district results using sound levels: The questions asked of participants concerning the election results were answered by some respondents based on auditory information and by some respondents based on interpretations of the visual map. For example, participants were asked which of the election contests in Ottawa Centre and Argenteuil-Mirabel they thought had a closer result.[2] The textual description accompanying

[2]As shown in Table B.2, the Bloc Québécois won the election in Argenteuil-Mirabel with 57% of the vote. The New Democrats won the election in Ottawa Centre with only 41% of the vote.

the election map states both that "[t]he speech of each party's leader is adjusted according to the party's percentage of the vote in that riding" and a user should "**move the cursor over a new district** to adjust playback according to the election results in that riding." Although this does not state that the gain settings associated with each speech are the sound parameters modified by the vote percentage variables, and being an earlier design this map did not include the dynamic audio-visual legends, the text, if read, would be expected to guide users to consider that sound playback is being adjusted based on the data.

Four of six (4/6) participants attempting to compare election results between the two districts independently decided, without assistance from the interviewer, to answer the question using the sound design. By contrast, P1 interpreted information in the visual map as providing a basis to answer the question and P8 first answered by interpreting information in the visual map but was then prompted by the interviewer to consider the sounds.

Three participants, P3, P5, P6, used the adjusted gain settings of the leaders' speeches to determine that Ottawa Centre had the closer election result. Two participants, after listening to the speeches in attempting to answer the question, decided that they could not answer.

The participants who attempted unsuccessfully to answer using the sounds both listened to the audio, moving the cursor back and forth between the identified electoral districts as they listened. P4, while initially studying the map, could be heard to read the description of the use of sound on the map and, when asked about the sound, eventually and hesitatingly[3] stated that he thought "the loudness of the voice is depending on who won or who came in what position [sic]." But when asked to compare the results in Argenteuil-Mirabel and Ottawa Centre, he

[3] The transcript of P4's answer to the interviewer's question "how many sound are there and how do they relate to the visual map?" follows (sic; interviewer reminders to answer inserted in bracketed italics):

> OK, when I click on a region, I can hear all the [pause] all or some parties speak at the same time. And it's hard to pay attention to any individual one [pause] and like the map is ... [pause] You want me to relate the sounds to the map? *[Yeah.]* OK, well, the map is pretty much like easy to see and it's organized because it's like the blues on the bottom and the reds on top and the colours are easy to distinguish, but when you listen to the sound it's hard to differentiate between the parties and pay attention to one single speech. *[So, how would you relate those people talking to the visuals? Like, do you think that they're related in any way or ... ?]* Um, OK. Well, I think like that when you go over the certain region, [pause] like when you go over the blue, then the Conservative, like the volume of his voice is higher. So you can't really hear the New Democratic voice on any of them. It's kind of faint in all of them, except for 62 where [pause] you can hear it, but whereas the other ones it's really hard to tell. I think, like, the loudness of the voice is depending on like who won or who came in what position.

During his answer, the cursor continued to flicker over visual elements without staying on any audio triggering element for long.

focused on the prevalence of the use of French in the former (i.e., the speeches of Gilles Duceppe, the leader of the Bloc Québécois) and English in the latter, constituted of the general mix of other party leaders speaking English with none predominating. Despite P4's stated belief concerning the use of volume to encode party standings in the vote results, he said that he could not decide which riding had a closer election result.

P8, by contrast, never understood the overlapped speeches as related to the data and, when asked, made the seemingly self-contradictory statement that the speech heard was "the specific party that won in that riding, but [...] it is confusing cause there is [sic] other voices that are playing in the background, I think maybe from the other [parties]." She seemed to pick out the loudest voice as related to the data but perhaps assumed that the others were not supposed to remain audible. She was not given assistance to understand the design and, when asked to compare the electoral riding results instead decided, based on the visual map, that the results for Ottawa Centre must have been closer because "it wasn't a very big riding." When prompted to consider the sound design in answering the question, she said she couldn't and subsequently, in response to an interviewer question, stated that she had not considered the sound volume levels as being related to the election results.

(ad.5) Only one of six (1/6) participants used information from a recorded speech to answer a question for which only that speech contained a particular perspective on the subject: The audio loop containing recorded segments of Paul Martin, the former Liberal Prime Minister, speaking on the night of the election begins with the statement "We as Liberals have lost votes, we have lost good members of parliament ..." During each evaluation session for OAFESM, to determine whether or not the participant was attending to narrative information in the speeches, the interviewer asked if they thought the Liberals would have considered the election to be a success. Only one participant, P8, said that the Liberals would not have considered the vote to be a success and explicitly referred to the recorded comments of Paul Martin. All other participants responded that, at least for the mapped region, the Liberals would probably be content because, as P3 said, "the majority of the map is red."[4]

As argued above, the participants all listened to the recorded audio (ad.2 on p.323). Therefore, the issue is probably not that most of them did not hear Paul Martin's statement concerning the disappointing results achieved in the election. Also, interpreting the phrase to be a self-referential opinion on the election outcome from the Liberal's perspective should have been straightforward because the recorded speech phrase explicitly identified the speaker as a member of the party:

[4]P5 asked whether or not the question referred to the election results overall and was told to assume "that we're talking about just the Ottawa [area]." This would probably have suggested that the visual map was the appropriate basis for his response. For the other participants, there appears to have been no such biasing guidance given.

"We as Liberals ..."[5] Quite possibly, the recorded statement having been played but not necessarily playing as the question was asked was less likely to be used as the basis of a response than the still visible categorical map of the Ottawa region results.

As discussed (ad.2 on p.323), some participants expressed concerns about not being able to isolate and listen to a single party leader's speech. Therefore, it is possible that a participant would not feel confident in basing an answer on the statements heard in a speech if they felt they may not have heard it clearly.

E.2.2 Canada's Trade with World Regions: 1976–2000 (CTWR)

(ad.6) Only two of six (2/6) participants decided without assistance that the music varied based on trade data values: Most users working with CTWR detected the playback changes in the music as they worked with the map, the dynamic audiovisual legends, or both. However, only two of six (2/6) decided, without prompting, that the acoustic changes related to changes in the values of the spatial variables. Instead, the evaluation session transcripts show that some participants initially related the visual changes of the dynamic audiovisual legends to the spatial data but did not infer a similar relation between the sound design and the data.

P1 and P8 both independently related the sound design to the trade variables. However, all other participants, after initially examining the application and verbalizing their impressions of the represented trade data, upon being asked about the sound design, responded that they did not understand it. These participants understood the representational role of the sound design only after the interviewer asked them to work with the dynamic audiovisual legends and describe the sound changes associated with the different data classifications.

The participants who did not understand the representational sound design ascribed other purposes to the use of sound. P2 discovered that the dynamic audiovisual legends are interactive controls before she discovered that the sounds are also associated with the map. Despite having spent time examining each visual meter, and hearing the acoustic changes associated with the classification range for each variable, once she was working with the map she seemed unsure that sound playback was altered based on the value of mapped variables, and asked "is the music just there for, like, soothingness [sic]?" P3 and P6 each asked if the music was "cultural," although P6 discounted this idea after determining that the same music was used for all regions. P7 expressed disapproval with the music when first asked about it, describing it as "distracting" because it seemed "happy" and unrelated to the data.

[5]The version of the map evaluated required that the user click on the map to start the audio looping and Paul Martin's statement, quoted above, is at the beginning of the Liberal audio clip, making it unlikely that this statement was not heard.

(ad.7) Five of six (5/6) participants were able to understand the sound design using dynamic audiovisual legends if they inferred a relationship between the visual meters and music: During the evaluation sessions for CTWR, the most common circumstances in which the interviewer provided assistance to a participant was when the participant expressed that they were not relating the sound design to the trade variables. In these cases, the interviewer asked the participant to use the dynamic audiovisual legends as interactive controls and to describe the changes they heard as they moved the cursor over the different classifications.

As discussed above, two participants, P1 and P8, independently used the dynamic audiovisual legends as interactive controls and, while moving the cursor over the different classifications, verbalized the sound changes they heard (ad.6 on p.328). However, other participants either did not use the legends as interactive controls until prompted by the interviewer or used the legends as interactive controls but did not notice the acoustic changes until so prompted.

For example, P1, having already examined the dynamic audiovisual legend associated with the *balance of trade* variable, began to move the cursor over the legend for the *value of exchange* variable and to verbalize the acoustic changes, saying "Oh, and apparently if I highlight the different [...] 'medium', 'low', and 'high' it makes it louder."

By contrast, P3 upon being shown CTWR and beginning to work with it, immediately started to move the cursor over the map and discussed the visual updating of the dynamic audiovisual legends with the interviewer, saying "it shows you current [...] net imports [...] ratio of Canadian exports to imports." Although she mentioned hearing a difference in the music when she moved the cursor from over the United States to the Mexico, Central America, and Caribbean region, she described it simply as "a different kind of music." When subsequently asked by the interviewer if she was noticing the musical changes, she indicated that she was but asked "is it like a cultural kind of thing?" The interviewer then suggested that she move the cursor over the dynamic audiovisual legend, whereupon P3 did so and began to verbalize the acoustic changes she was hearing. For example, with the *value of exchange* meter, she stated "So, it gets louder and there's more instruments in it when you get [sic] from low to high."

Although P2 used the dynamic audiovisual legends in CTWR extensively as interactive controllers during her evaluation, she was initially unclear concerning the relationship between the music and the mapped variables, as discussed above (ad.6 on p.328). Once P2 asked about the purpose of the sound and questioned whether or not it was intended to sooth the user, the interviewer responded by providing additional assistance, directing P2 to place the cursor over the visual meters again and asking her to describe the acoustic changes associated with each classification of the spatial variables as they worked together with the meters. Once this exercise was completed, P2 successfully answered questions using the auditory representation with the visual meters hidden.

Five of six (5/6) participants were able to verbalize the changes they heard in the sound design while working with the dynamic audiovisual legends, either independently or upon being prompted to by the interviewer. Once they had completed this exercise, they were able to distinguish the sounds associated with different value ranges while working with the map. But there was a remarkable tendency among participants not to consider the sound design as being related to the mapped data until such a relationship was suggested by an interviewer.

(ad.8) Participants used the visual design when learning the sound design, and two of six (2/6) participants continued to use visuals when interpreting mapped data: As discussed above, interacting with the dynamic audiovisual legends while listening to the acoustic changes triggered by that interaction seems to have been a necessary step in understanding the sound design for CTWR (ad.7 on p.329). Even after having understood the sound design, participants can be heard to describe it, when asked, by referring to the dynamic audiovisual legends and how the sounds function in relation to these visual interface elements. P1 who independently used the legends to learn the sound design, when responding to the interviewer's request to explain the relationships between sounds and the map, described the sound design for the *value of exchange* variable in relation to the visual design by saying "[...] there's two [sic] different columns [dynamic audiovisual legends], one column indicates the percentage, it's 'low', 'medium', or 'high', and it's the same sound but it gets louder as it gets lower or higher [...] if it's low, it's not as loud and if it's a high percentage it's louder." P1 then continued on to describe the sound design for the *trade balance* variable using similarly visual references to the dynamic audiovisual legend for that variable. During the evaluation session, P1 responded to this question by moving the cursor over the different classifications on the legends as she described the effects of the changes in variable value, suggesting that interaction with the legend provided a useful reminder for the participant in the act of describing the design.

After having learned the sound design for CTWR, five of six (5/6) participants successfully answered at least one question concerning Canada's trade with world regions using the sound design when the dynamic audiovisual legends were not displayed. But when offered the option to use the visual design, three of six (3/6) preferred to do so. P2, when asked about the classification of Canada's trade balance with world regions, asked if she was expected to answer "just by the sound." After being told that she could use any means she wanted, she used the visual legends, explaining that "I'd use the words and the chart and stuff to figure it out because [...] I don't link the sound to the numbers." Subsequently, P2 was asked to answer questions without the use of the dynamic audiovisual legends and she successfully completed those tasks.

Two participants, P3 and P6, when asked to answer questions without the dynamic audiovisual legends being visible, preferred to and were allowed to answer at least one of the questions with the legends redisplayed. For P3, this may indicate a preference rather than an inability to answer without the visual design because she did answer one of the questions using only the sound design. By contrast, P6 was confused by at least the trade balance audio indicator and was never able to verbalize how the instruments for that indicator were being controlled. Therefore, his switching back to use the visual legends is almost certainly a further indication that he was having difficulty with the sound design.

E.2.3 Antarctic Territorial Claims (ATC)

(ad.9) Four of six (4/6) participants disliked that they did not know how to isolate each speech: The overlapping map sectors of ATC enable a map user to isolate the Chilean and British news narrations but there is no way using the map to isolate the Argentinian narration because the entire Argentinian sector overlaps one or both of the British and Chilean sectors. Because the country names in the text panel are also hyperlinks that adjust the gain setting for the associated looping news narrations, a user can place their cursor over a country name to isolate the narration of a newspaper story related to one of the national claims. The textual description accompanying the map does not mention that the country name hyperlinks initiate narration playback.

During the evaluation sessions for ATC, four of six (4/6) participants initiated playback of the claim-associated narrations by moving the cursor over the map resulting in multiple narrations starting to play simultaneously. These participants all expressed concerns about the lack of ease in being able to isolate a single narration or focus satisfactorily on what was said in each narration. Through working with the map, P2, P3, P4, and P5 all initiated playback of multiple narrations simultaneously. Having decided that the male English narrator was translating what was said by the female Argentinian speaker, P2 commented that the female voice "is louder to me so I find it really hard to understand what the guy is saying." Subsequently, in response to an interviewer question, P2 clarified that "I'm just trying to understand what they're saying. [...] I want to get the content but right now it sounds like there's three or four voices talking at once and it's so confusing."

P3, upon first hearing the overlapped voices while interacting with the map, moved the cursor slowly across the contested sectors and listened to the varying resultant narration mixes before stating "[t]hat got really confusing for a while."

P4, similarly, when hearing the blended voices while moving the cursor over the overlapped territorial claims, said "So all these people came up at once and started talking. [...] I definitely don't like the fact that everything comes up at once."

P5 could be heard to react negatively to the blended narrations on the map and responded with satisfaction when he isolated the English narration. "[... Cursor over UK sector ... cursor over UK/Argentina sector ... cursor over UK/Argentina/Chile sector ... cursor moves off map ...] They can overlap? Wow, that's weird. [... cursor over UK, only UK voice is heard ...] There we go." In response to a subsequent interviewer question concerning whether or not he had been attempting to isolate a single voice, P5 responded that "if I found this online like this, I'd consider that to be a bug."

Both P6 and P7, by contrast, began their use of ATC by reading sections of the text panel. In the process, both first initiated playback of the territorial claim narrations through the use of text hyperlinks, thereby isolating individual narrations. Each listened to segments of the different narrations before continuing on to examine the map, whereupon each eventually initiated simultaneous playback of multiple narrations through the map's overlapping sector design. Neither voiced concern over the blended playback of narrations.

(ad.10) Three of six (3/6) participants, when asked, stated that the recorded speeches indicated the overlapping national claims to territory: P4, P6, and P7, when asked by the interviewer, said that the overlapped speeches showed the different claims or 'conflicts' concerning territory in Antarctica. The other three participants did not relate the overlapped speech design to the territorial claims theme of the map. P2 decided that the multiple recorded speeches were alternative language translations of the same story concerning the territorial claims of Chile, Argentina, and the United Kingdom. When asked why only some country claims would have associated sounds, she guessed that only those three countries were involved in a dispute but continued that she didn't know why the map didn't provide verbal narration for the other non-disputed claims as well.

P3 stated that she thought the sounds were "just people talking different languages" and they're "the languages of the specific region." When asked why some sectors had associated sounds and others didn't, she said she did not know.

P5 initially stated that he thought he would interpret the overlapping narrations design as a defect if he found the map online. But he continued that he thought the narrations were national perspectives on the territorial claims, correctly guessing that the Spanish speakers were also reading newspaper stories related to the claims, similar to that being read by the British speaker, whose narration he isolated and to which he listened in full. However, when asked why other map sectors had no associated narrations, he said he didn't know and guessed that perhaps the map was still being developed and was incomplete.

These participants' statements were made despite the text panel initially showed to them stating that "[t]he function of vocal narration [in this application] is to highlight the tension between these three countries [United Kingdom, Chile, and Argentina]." This statement, if read, would be expected to clearly indicate that sound should be expected over only some of the claims, that there was a purpose behind the design applying sound use only on some of the mapped claims, and moreover what that purpose was.

(a) (b)

Figure E.1: Cybercartographic Atlas of Antarctica: Text panels for contested sectors: a) claimed by Argentina and the United Kingdom; b) claimed by Argentina, Chile, and the United Kingdom (image produced by Glenn Brauen using screen captures courtesy of the Geomatics and Cartographic Research Centre, Carleton University).

(ad.11) The mixed narrations design reinforced contested claims but did not independently signify the concept: As discussed above, P2 and P3 did not interpret the mixed narrations design of ATC as signifying the national contestation of Antarctic territory (ad.10 on p.332). By contrast, P5 tentatively suggested this meaning but contradicted that idea by suggesting that the other map sectors did not have sound because the map was simply not finished. P5, upon beginning to work with ATC, appeared to read the text panel, scrolling through it and pausing, before he began to move his cursor over the map. Therefore, since the playback of narrations as a result of moving his cursor over the map did not clearly suggest the contestation of national claims, his interpretation that this was the theme of the map is as likely to have resulted from one or more of the displayed atlas section title "Sovereignty related Conflicts" (see Figure B.5), the interviewer's introduction of the map as being about "Antarctic territorial conflicts," or information in the text panel.

P4, when asked to describe the sounds and relate them to the visuals, proceeded to read the currently displayed text panel, describing the sector contested by Britain and Argentina (see Figure E.1a).

> So [...] this slice of Antarctica is claimed by the United Kingdom and by Argentina. Historically the UK was the first nation to claim a part of Antarctica. Argentina claim [sic] was made three decades later [... reading becomes inaudible ...] I think the sound has to do with the conflicts. So they both talk at the same time so it's kind of symbolizing like [sic] conflicts of their speech.

As P4 reads the passage from the text panel, the cursor remains over the map sector, causing the British and Argentinian narrations to play while he reads. It is only after reading the text describing differing national claims to the map sector on which he had most recently clicked that he answers that "the sound has to do with the conflicts." But his answer also specifically refers to the blended narration and the symbolism of their simultaneous playback, suggesting that this did reinforce the contestation for P4.

While examining ATC, as discussed above (ad.10 on p.332), P6 began by reading the text panel. He continued to pay attention to the text panel as he eventually started to move the cursor over the map and initiate blended playback of the narrations. Just before he began to verbalize the behaviour of the narrations in relation to where his cursor was positioned, he clicked on a series of the map sectors and discussed the updates he saw in the text panel as a result, finally leaving the panel displaying the image and text description of the map sector claimed by all of Argentina, Chile, and the United Kingdom (see Figure E.1b). Thereafter, he could be heard to paraphrase the written chronology of these nation's claims before discussing the playback of the narrations.

> Historically, Britain was the nation to claim that part of it but then Chile and Argentina claimed it later. [... cursor over New Zealand sector ...] Um, [... cursor over United Kingdom/Argentina/Chile sector ...] so [... cursor over United Kingdom/Chile sector ... cursor over Chile sector ...] this area perhaps seems to be only one voice, so maybe there is no dispute in that area. [... cursor over United Kingdom/Chile sector ...] Two voices start up here so perhaps two countries and actually it is indicating Chile and United Kingdom have stakes in this area. [... cursor over United Kingdom/Argentina/Chile sector ...] Three voices now start and Argentina has lit up on the side so all three here. [... cursor over United Kingdom/Argentina sector ...] All of them except for Chile here. [... cursor over United Kingdom sector ...] And here just United Kingdom. [... cursor off map ...]

Throughout the above quote, P6 continued to refer to the visual updates in the text panel as hyperlinks "lit up" because of his cursor movements on the map, making it clear that he was attending to visual updates resulting from his actions and the narrations he could be heard enumerating.

P7, similarly to P6, began his examination of ATC by reading the text panel and began to discuss the contested claims in relation to the text hyperlinks he referred to as "the different highlighted blue items" before he began to move the cursor over the map and initiated mixed narration playback. While moving the cursor over the map, highlighting both contested and uncontested sectors and thus initiating narration playback intermittently, he described the contested national claims in primarily visual terms that, in fact, make it uncertain whether or not he was attending to the narrations as he worked with the map.

> [... cursor over New Zealand sector ... cursor over Australian sector

> ...] So I guess this is just [... cursor over Norway sector ... cursor over United Kingdom/Argentina/Chile sector ...] a big pie section of who owns what. [... cursor over United States sector ... cursor over United Kingdom/Chile sector ... cursor over United Kingdom/Argentina/Chile sector ...] And this looks like an overlapped dispute. [... cursor over United States sector ...]

However, as discussed above, P7's first response to hearing the narrations had been to comment on the emotional significance of the narrations, describing a speaker as "excited" and the general tone of the narrations as "heated" (ad.12 on p.335). Therefore, it is reasonable to assume that, despite his primarily visual response to the immediate question concerning the application's sound design, the mixed narrations did at least reinforce for him the potential for dispute concerning the mapped national claims.

(ad.12) Four of six (4/6) participants listened to and commented on the languages used in the recorded speeches: Upon hearing the use of overlapped speeches and noting that one was speaking a language other than English, P2 remarked that the English speaker was probably a translation of what was being said by the other. P3, P4, and P5 decided that language was being used to signify different national claims to Antarctic territory with an English speaker voicing the claim of the United Kingdom while Spanish speakers represented the Argentinian and Chilean claims. In addition, P5 remarked that he would have preferred that the map include a translation of the recorded Spanish in some form because he was "sure it's not just from the British point of view."

In this respect, P6 and P7 were exceptions and did not comment on the languages used in the speech. P6 remarked only on the presence of voices in relation to the disputed sectors on the visual map. P7 commented on the emotional significance of the speeches, noting that the female Argentinian speaker seemed "excited" and that the overlapped speeches indicated a dispute "because they seem very heated."

E.2.4 Antarctic Exploration (AE)

(ad.13) Five of six (5/6) participants detected qualitative changes in sound effects without being prompted: As discussed in Section B.3.2, the evaluation sessions for AE all began with the interviewer asking the participant to invoke certain features of the application, one being a chronological display of selected expeditions initiated by the user sweeping the cursor from left to right across the array of expedition selector buttons. The resulting display sequence mimicked an animation of the chronology of expeditions and showed a series of expedition route maps in which the "Myth of the Island of Utopia" was "whittled away" by the incremental understanding of continental Antarctica. During the chronology, each route map was accompanied by audio indicating the predominant form of transportation used by the expedition.

Five of six (5/6) participants listened to and commented on the succession of sounds associated with the Antarctic expedition chronology. With minor variations, all of P2, P4, P5, P7, and P8 discussed the sounds heard and interpreted the sounds as representing modes of transportation used by the currently selected expedition. The transcript of P2's evaluation of this part of the application is typical of those of this group of participants.

> [... scrolling cursor across early explorers on timeline ...] I don't know what this means yet. [...] The map in the middle keeps getting blurrier every time I click across it. [... cursor over Shackleton, sound shifts ...] Oooh, it's windy there. [... cursor over Amundsen ...] So, I'm guessing that the wind indicates that they're actually on the island now, versus just... [... cursor over Byrd, wind sounds subside, leaving momentary silence before plane is clearly audible ...] Oh, now he's fallen over? [... airplane becomes audible ...] Sounds like an airplane. Sounded like sea kind of stuff, and the wind they actually got on to it, and now something about airplanes.

Not all participants correctly identified each of the sounds, and some subsequently used available visual or textual information to confirm or correct their first impressions of the sounds. In the above section of the transcript for P2's session, she does not explicitly mention the sound of footsteps in the snow as part of the sound accompanying the Shackleton and Amundsen expedition displays, although her suggestion that "they're actually on the island now" could have been cued by the footsteps.

P5, by contrast, did discuss the sound of overland trekking when these expeditions were selected, saying "It sounds like a bitter wind whipping across an arctic tundra, or Antarctica I suppose in this case, and somebody walking, some poor masochistic soul [...]"

P8 stated that she heard the sound of someone walking in the audio associated with Shackleton's expedition but P4 simply described the sound, in terms not dissimilar to those used by P2, as being "gusts of wind blowing through something."

P7, having finished displaying the chronology of expeditions, discussed all of the transportation modes used by the expeditions except for overland trekking and subsequently searched through the sounds again until he found the audio associated with the Shackleton expedition. He then discussed the sound, stating that interpreting the sound as someone walking to get to Antarctica was unreasonable.

> Sounds like they walked it. [... cursor circling red route line and footprint icons on the map ... cursor over text panel map and then circling the area of open water between South America and the Antarctic Penin-

> sula ...] I don't think there is this continental shelf that allows them to walk it.

Whereas all other participants working with AE commented on the sounds associated with the expedition displays without needing any prompting from the interviewer, P1 commented almost solely on visual components of the display until prompted by the interviewer to attend to the sounds. For example, among other things, she discussed the visual progression showing "the formation of Antarctica," expressed approval for the route maps, and discussed the tendency of the multimedia panel accompanying the map to display less information for the later expeditions. Her only comment concerning the application sound design, prior to the interviewer's prompt and at a time when she had only heard the sound associated with the seafaring expeditions, was her statement that she was "[n]ot big on the waves though." During her display of the expedition chronology, the sound changes are clearly audible but she can be heard to discuss only the visual changes, as above. Once the interviewer prompted her to listen for the sounds, she redisplayed the expedition chronology and commented on the sounds in similar terms to those used by the other participants. Therefore, she was able to hear the changes but for some reason did not pay attention to the changing sound effects when first working with the expedition chronology.

(ad.14) Five of six (5/6) participants recognized the fly-over sound effect as an airplane: All participants recognized the sound associated with the fly-over expeditions as being the sound of a small airplane except for P8. She misidentified the sound as that of an helicopter.

(ad.15) Five of six (5/6) participants interpreted the NASA sound effect as 'space sounds' or RADAR but one had to use visuals to reach this conclusion: Four of six (4/6) participants interpreted the sound associated with the National Aeronautics and Space Administration (NASA) imaging of Antarctica as the sound of satellite imaging, RADAR, or 'space sounds' during the expedition chronology (i.e., when no visual images were shown to clarify the sound). In addition, P4 had initially asked if the sound was that of a submarine but clicked on the NASA expedition selector after completing the sweep of the expedition chronology and used the map icons and the images in the text panel to conclude that it was the sound of a satellite. P1, upon hearing the NASA-associated sound, stated "I have no idea what that is!"

(ad.16) None of the participants (0/6) related the map background transformation to the "Island of Utopia" narration to which they listened: As discussed in Section B.3.2, when beginning their evaluation of AE, participants were asked to first listen to a narration that very briefly discussed the myth of a great southern island and how the history of exploration had whittled down that myth to produce contemporary understandings of the continent. Participants would then be asked to sweep the cursor across the expedition selectors array to cause a chronology of expeditions to be displayed, and to comment on what they thought the sounds meant.

While displaying and discussing the sounds and the expedition route displays, none of the participants related the map background updates, structured as a transformation of "The Island of Utopia" into a contemporary relief image of the continent, to the narration to which they had just listened.[6] Three participants, P1, P2, and P4, commented on the increasing blurriness of the map background as they progressed through the middle of the expedition chronology and P1 commented that maybe the map was "showing the formation of Antarctica."[7]

Participants may not have related the narration and the imagery transformation because of the time delay between hearing the narration and their beginning to display the expedition chronology, although the span of time was fairly brief. More likely, the interviewer immediately following the "Island of Utopia" perspective on the atlas section with new instructions and the added interruption of the expedition chronology sound effects distracted the participants from that perspective and they never reconsidered it. While examining the expedition chronology and because of interviewer prompting to explain what they thought the sounds indicated, all participants discussed the simultaneously audible sound effects in relation to the visual map updates showing the changing background and the expedition routes.

(ad.17) Participants interpreted information using both sound and visuals: Either when examining the expedition chronology or when answering interviewer questions concerning the modes of transportation used in exploring Antarctica, several participants used both the sound design and visual information available in AE to formulate and confirm their interpretations and responses. P1, P5, and P7 all interpreted the sound of wind-whipped footsteps in the snow as indicating that the expeditions had gone overland and each of these participants confirmed that interpretation using visual iconography and textual explanations displayed when the associated expedition selectors were clicked. When previously asked what was the predominant mode of transportation used to explore Antarctica, P5 had answered by sweeping the cursor from left to right across the selectors until he

[6]The backgrounds are updated with each expedition selection and show a timeline indexed visual transformation that begins with an illustration titled "The Island of Utopia" and ends with a terrain elevation rendering of Antarctica. The "Island" illustration is a reproduction of *Utopiae Insulae Tabula*, a woodcut by Ambrosius Holbein originally published in the 3rd edition of Sir Thomas More's *Utopia* (1518, 12).

[7]Despite the script for evaluations of AE, P1 was not asked to sweep the cursor across the expedition selectors array immediately after listening to the "Island" narration and was instead told to simply explore the application. Therefore, she spent more time looking at the application, focused mainly on the visual updates and the contents of the text panel accompanying the map (ad.13 on p.335) and the time lag between hearing the narration and seeing the map background transformation as part of the expedition chronology was greater than for other participants.

found the Shackleton expedition, that being the first not associated with the sound of water and waves. Then when asked who first explored Antarctica on foot, he immediately answered that it was Shackleton but then confirmed his answer by clicking on the Shackleton selector, examining the text panel, and then explained "There are little footprints on the map which sort of indicate that he's wandering around on there [...]" Both P1 and P7 similarly answered the question concerning the first overland expedition first by using the sounds and then confirmed their interpretations using visual information.[8]

Upon being asked what she thought the sounds triggered by the expedition chronology represented, P2 described them in terms of the modes of transportation she thought they represented but expressed uncertainty concerning the relevance of the drone of an airplane. To clarify that, she redisplayed the expedition chronology and used the sound along with the changing map imagery to interpret the meaning.

> I'm not quite sure what flying has to do with it. [... cursor moving over expedition selectors until over Byrd, airplane sound resumes ...] I don't know, maybe they were taking pictures of it. That kind of looks like they're developing the island, what it actually looks like?

As discussed above (ad.13 on p.335), P4 initially suggested that the sound associated with the NASA imaging of Antarctica was that of a submarine. After having completed the display of the expedition chronology as suggested by the interviewer, his first action was to click on the NASA expedition selector, causing the map display and multimedia panel to be updated to include satellite icons and images. Having seen these, P4 then asked "Was that the sound of a satellite?"

When asked what was the predominant mode of transportation used to explore Antarctica, P8 answered that it was most often done by sea voyage and, when asked how she got that answer, explained "there's a lot of storms [...] so there would be a lot of boats." She subsequently confirmed this answer by referring specifically to the voyage routes and ship icons displayed on the expedition maps in the atlas section, saying "most of them would go by boat because most of those pictures all have a boat route."

(ad.18) The transcripts show two (2) participants interpreting information using sound without acknowledging that source of information: The evaluation session transcripts contain instances in which participants commented on AE or responded to interviewer questions after working with the application in ways that strongly suggest they were attending to the sound design. In two of these

[8]According to the displayed route maps, the earliest overland expeditions were actually conducted by Robert Scott, whose expedition is represented on the route map as a combined seafaring and overland expedition and in the audio using the seafaring sound loop. Only P2 answered that Scott's expedition was the first to explore Antarctica by foot based on that visual route map.

cases, the participants subsequently answered interviewer questions concerning how they got that information by emphasizing sources other than the sound design. It is unclear that their interpretation would have been sustained without having listened to the sound design or was not at least partially influenced by that in addition to the visual information to which they explicitly referred. This may indicate a lack of awareness concerning their response to the sound design.

Having completed his examination of the expedition chronology, P4 stated that he would like to "[l]ook at some pictures and get an idea of what's going on." Whereupon, he clicked on the Amundsen expedition selector, briefly causing the sound of overland trekking to resume, and stated "This one's probably the sound of wind because [...] when I look at this picture I think of like the explorer on his sled with dogs as opposed to like a boat or a plane [sic]. And [...] right away from the image I can get [...] the sound of wind." P4 subsequently confirmed it by replaying the audio associated with the Amundsen expedition, commenting "Yeah, so I was right, it is the sound of wind." However, when asked how he had confirmed that the sound was wind, P4 said that the picture confirmed for him that the sound was wind.

The picture referred to by P4 (Figure E.2), is a photograph of Oscar Wisting, a member of Amundsen's expedition, standing by a sledge and in front of a Norwegian flag planted at the pole with a dog team lying on the snow in the foreground.[9] Although the flag is fully extended, that is the only feature of this image that would suggest the characterization 'windy.' The packed and clearly tracked snow around Wisting and the dog team coupled with the relatively clear visibility of the image background in which their tracks can be seen to recede into the distance, without the presence of blowing snow, indicate that, when the photo was taken, winds were light although strong enough to support the flag. Rather than suggesting windy conditions, this image, if presented to me, would suggest any of 'cold,' 'barren,' or 'isolated' more strongly. P4's interpretation of the visibility of the wind in the image could have been suggested by the extended flag alone but would more convincingly have been at least partially in response to the sound of overland trekking, replayed as he worked with the timeline to display the details of the Amundsen expedition just prior to his comments, and subsequently reinforced by the image and listening again to the audio associated with the expedition.

As discussed above (ad.17 on p.338), P8's response concerning the predominant mode of transportation used by Antarctic expeditions was strongly influenced by the accompanying sound of wind and waves, or 'storms' as P8 referred to them. Although she subsequently confirmed the predominance of seafaring as a mode of exploration using the visual route maps and ship icons, these map displays do not suggest storms. The ship icons show a cartoon graphic of a single mast sailboat in relatively calm seas and the maps depict the southern oceans as a uniformly placid

[9]The image was first published by Amundsen (1912, 120), although the atlas used a copy of the image from McGonigal and Woodworth (2001, 451).

Figure E.2: Cybercartographic Atlas of Antarctica: Oscar Wisting and dog team at the south pole. Reproduced from Amundsen (1912, 120), now in public domain.

blue surface (see Figure B.6). The seafaring audio loop does not strongly suggest storms but the wind and wave events depicted in the sound design would more readily carry this interpretation than do the route map displays. When asked by the interviewer whether or not she had used "audio at all to get that information," P8 responded that she had not.

Although she confirmed her response concerning the predominance of seafaring expeditions using the visual maps, P8's first response that "there's a lot of storms [...] so there would be a lot of boats" contains a reference to weather conditions that is more convincingly supported either by the sound design or by the sound design in combination with the visual map than it is by the visual information alone.

E.2.5 General

(ad.19) Some participants expressed surprise upon hearing sound: In some cases, participants working with the audiovisual map examples expressed surprise when their interactions with a map initiated sounds. This despite these students having been recruited to participate in these map evaluation sessions through the use of postings that advertised the study as "Usability Tests for Multimedia Maps that Use Sound" (Appendix D). Expressions of surprise were most often exclamations uttered by the participant upon hearing sound but in two of the cases shown

Table E.2: User evaluations: Expressions of surprise at use of sound, by map example: entry indicates the order in which the map eliciting surprise was presented to the participant (among the three maps presented to each). Greyed out cells show cases where a participant did not test a map example.

	Participant Surprised by Sound								**Totals,**
	P1	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**	**P8**	**by map**
Election								1st	1
Trade	2nd	1st						2nd	3
Claims			1st		2nd				2
Exploration									0

the interviewer can be heard to say that the participant looked surprised, whereupon the participant confirmed it.

As shown in Table E.2, five of eight (5/8) participants in these evaluations expressed surprise when working with at least one of the map examples and P8 expressed surprise while working with two of three map examples. Three of six (3/6) participants working with CTWR expressed surprise when they first moved the cursor over the dynamic audiovisual legends. Two of six (2/6) participants working with ATC expressed surprise when they first moved the cursor over the contested sectors of the map. One participant of six (1/6) evaluating OAFESM expressed surprise just after moving the cursor over one of the electoral districts on the map.

It is difficult to draw conclusions from the fact that these participants expressed surprise at the occurrence of sound, but two speculations may be proposed, each of which requires further study. First, acclimatization may have already occurred for most participants during the time they were working with the audiovisual map examples. All expressions of surprise occurred during their evaluations of the first and second example with which they were presented and no participants expressed surprise while working with the third of the map examples they evaluated. If acclimatization to the idea of using sound in maps was occurring, this may indicate that the lack of consideration of sound in cartographic practice extends beyond the population of map researchers and map-making practitioners.[10]

Second, no participant expressed surprise when sound effects playback was initiated within AE, regardless of where in the order of evaluations conducted with that participant this example occurred (see Table 6.1). This may indicate that participants were expressing surprise as a result of the particular ways in which sound was used rather than because of the mere presence of sound. As discussed in relation to the apparently greater propensity of these participants to consider the sounds and visual updates of AE as semantically related (Section 6.3), a greater fa-

[10]See Section 1.3 concerning the coverage of sound in the cartographic research literature and a sampling of maps on the World Wide Web using sound in some way.

miliarity with the use of sound effects in conjunction with visual media compared to the novelty of other sound designs used in these examples (e.g., overlapping speech) may explain this discrepancy.

Appendix F: Sound Subsystem Software Overview

F.1 Brief History and Acknowledgement

The original versions of this software were developed by the author as part of this thesis research and as part of his work as a Research Assistant with Carleton University's Geomatics and Cartographic Research Centre (GCRC) during 2004 and 2005. The software was implemented as a Java applet providing an Application Programming Interface (API) enabling WWW browser pages to load and control playback of sound media files. The software supported the playback of multiple recorded ('sampled') audio tracks including separate control of playback start, pause, and stop for each track and separate controls to perform *fading* gain adjustments, incremental adjustments over time to avoid abrupt and harsh acoustic events, for each track.

Jean-Pierre Fiset, a colleague working with the GCRC, suggested that the responsiveness of interactive maps created using the sound applet would be improved if we created a control library that decoupled the handling of user interface event tracking[1] from the actual work of changing the current sound playback parameters. He then created a Javascript library, the *sound proxy*, that collected parameter updates for the sound applet but only sent these updates to the sound applet as a background batch activity, thereby significantly improving the responsiveness of maps created using the sound applet.

In 2006 and 2007, J. P. Fiset added support in the sound applet to enable the sound applet to manage a very large number of sampled audio files being loaded and unloaded over time as a result of user actions. The large number of files was required to use the sound applet with a place names pronunciation map with each audio file containing the pronunciation of the name of a single place on the map.

At about the same time, during 2006 and during the summer and autumn of 2007, the author added support to the sound applet and sound proxy to enable playback and control of Musical Instrument Digital Interface (MIDI) compositions, providing independent playback controls for gain and pitch transposition on each

[1]For example, detecting that a user has moved the cursor over a particular map region and determining the implications of that operation for the control of the visual and auditory feedback behaviours of the map.

instrument channel.

During 2009, the author added support for run-time effects processing of sampled audio tracks in both the sound applet and sound proxy and added Javascript libraries that implemented the modular model described in this thesis for audio indicators structured as sets of audio sources and effects processing chains. This work included Javascript software components for monitoring the currently selected value of spatial variables defined by the implementation of an interactive map and for using this value to compute updated values for sound parameters defined by the processing effects chain of an auditory indicator to update the sound proxy.

The media CD-ROM included with this thesis contains source code for all of the components described above. See Appendix G and the contents of the Media and Source Code CD-ROM for more details.

F.2 Sound Subsystem Software Components

Figure F.1 shows the major software components used to manage the auditory representations designed into an audiovisual web map application. Some of the main functional interfaces for each of the software components are shown although the diagram is incomplete. For example, the sound parameter functions shown for the sound applet and proxy are an extensively-used subset of the functions specific to managing sampled audio. An equivalent set of interfaces exist for use with MIDI compositions but are not shown here. Although Figure F.1 shows only functional interfaces related to managing sampled audio, the logic of the division of responsibilities within the software design is very similar for the two cases of MIDI and sampled audio. Similarly, in all software components, functional divisions between MIDI and sampled audio management are fairly clear because all interfaces involved in the management of MIDI composition playback include 'Midi' in their name (sampled audio was done first so MIDI became the later, exceptional case). This section is intended to provide a brief overview and enough information to assist a reader in determining what files to examine to more fully understand the software design for managing audiovisual map sound. The software for all of the components described here are included on the CD-ROM accompanying this thesis. Appendix G details the contents of the accompanying CD-ROM, including identifying the organization of the software, and identifies key files (i.e., good starting points) for each of these components.

Web Browser Application The audiovisual map application operates in the context of a web browser that interacts with one or more hypertext transfer protocol (HTTP) web servers, some of which also provide the Open Geospatial Consortium (OGC) map services (e.g., Web Map Services (*WMS*, see OGC 2002), Web Feature Services (*WFS*, see OGC 2005)). Although this component is shown as interacting

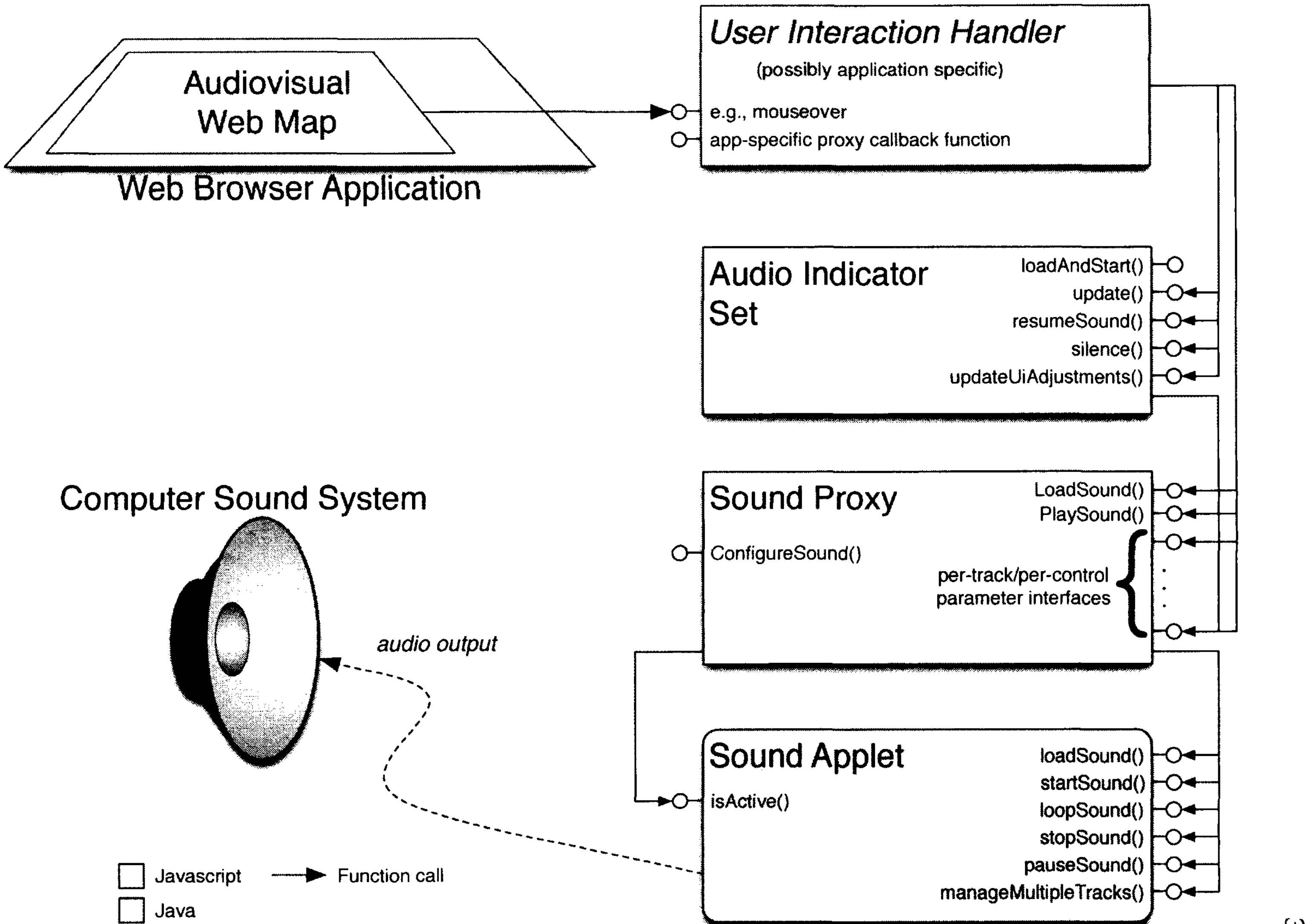


Figure F.1: Major sound subsystem software components and exemplar functional interfaces used for sound updates (figure concept and design: Glenn Brauen).

with the Javascript software components of this design, the web browser creates the run-time processes in which the Javascript components execute and creates the Java run-time environment in which the sound applet executes.

User Interaction Handler As a user works with an audiovisual map application, software handlers detect application events such as the cursor moving into, out of, or within map features. The events are handled in application specific ways, depending on how the map is designed to work. Some of these application events will cause updates to monitored variables that are associated with auditory representations. These updates will cause the map interaction handler to invoke functions on the audio indicator set designed for the application.

Audio Indicator Set The audio indicator set maps spatial variable updates to sound control parameter updates, for both sampled audio and MIDI instruments, as outlined in Chapter 7.[2] During normal operation of the map application, the following functional interfaces are most used:

- `update`(*key*, *valueObj*, *muteFlag*): the audio indicator identified by *key* is updated using the value defined by *valueObj*. The parameter *muteFlag* is a Boolean variable indicating whether or not the entire sound system has been muted, probably in response to a device such as an application mute button. The parameter *valueObj* is an object containing two attributes:
  - *defined*: a Boolean variable indicating that the value is valid, otherwise the value is undefined and should be ignored (for example, when the cursor has moved off the map). Audio processing components must define behaviour for handling both defined and undefined values.
  - *value*: a numeric value computed for the spatial variable represented by the audio indicator identified by *key*.
- `resumeSound`(*key*, *muteFlag*): the audio indicator identified by *key* should resume audible playback, subject to the setting of the carried *muteFlag*. This interface is used typically to restart playback of an indicator after it has been muted, along with all other audio indicators, by the `silence`() interface. The parameter *muteFlag* is a Boolean variable indicating whether or not the entire sound system has been muted, probably in response to a device such as an application mute button.
- `silence`(): the sound subsystem is silenced, usually in response to an event such as the cursor moving off the map. This interface mutes all sounds which

[2]The functions shown here for the audio indicator set are an exception to the separation of interfaces based on which of sampled audio or MIDI they deal with. The configuration of sound designs using audio indicator components, as shown in Chapter 7, do differentiate between sampled audio and MIDI components but an auditory indicator design could validly contain both types of audio components. The audio indicator set functions responsible for updating and managing the running audio designs must then be able to deal with both sampled audio and MIDI components.

can create a rather sudden and harsh sound transition. An alternative to using this interface would be, for example, to define a special geometry around the main visual features of a map and define the event handler such that the cursor moving over this border geometry causes all indicators to be updated with *undefined* values. The use of the `update()` interface with an appropriately configured muting sound design (including a fade out timer longer than the default) can provide a less harsh sound adjustment.

- `updateUiAdjustments`(*key*, *nonDefaultSettings*, *defaultSettings*): the sound controls of the audio indicator identified by *key* are adjusted according to the difference between the default settings of the user interface controls for the indicator and the current interface settings, identified as the *nonDefaultSettings*. The user interface settings currently support gain and pan controls.

Sound Proxy The sound proxy collects audio control parameter updates, indicated through calls made to the per-track/per-control interfaces it provides for use either directly by the user interaction handler or by the audio indicator set, and asynchronously invokes sound applet interfaces to manage sound playback. Some of the operations of the sound applet can be slow when invoked (e.g., initiating the downloading of a sound file and configuring the audio format conversions required to play it). Thus the sound proxy improves the responsiveness of the audiovisual map by allowing the user interaction handler to quickly pass off updates without waiting around for the sound parameter changes implied by those updates to be finalized, allowing it to be quickly ready to deal with the next user interaction.

Sound Applet The sound applet is a Java applet embedded as an invisible component of the web application implementing an audiovisual map and manages all details of sound playback and control.

The sound applet software is divided functionally into two sub-components, the first managing the playback and acoustic parameter controls for loaded sampled audio files and the second managing the playback and parameter controls of all instruments in a single MIDI composition.

For sampled audio, each loaded track is identified by the Universal Resource Locator (URL) used to access and download the file containing the audio samples. Subsequent to it being initially loaded, all parameter updates for that sound entity include that URL as a key to identify the operational sound for which control parameters are to be updated.

For MIDI music, the applet supports only a single MIDI composition because the benefit of using MIDI is the maintenance of synchronization between a number of separately controllable musical instruments. The current Java software MIDI implementation can only guarantee synchronization within a single composition so there would be no benefit to trying to load and manage more than one (i.e., they would not be synchronized and unsynchronized sound can be accomplished with sampled audio). The parameter controls for MIDI playback supported by the sound applet are addressed to separate tracks and channels within a composition.[3]

F.2.1 Sound Subsystem Initialization

Having briefly outlined the processing interfaces used to update and manage a sound design once it has been configured, this section provides an overview of the functions used to configure a sound design during the initialization of an audiovisual map application. Figure F.2 shows the major software components of the sound subsystem and the role each plays during initialization of the sound subsystem. The timing of the creation and initialization of separate parts of a web application within the browser Document Object Model (DOM) is undefined and depends to some degree on browser and computer platform implementation details as well as network delays. This indeterminate timing means that the software design must dictate the ordering of dependencies in the initialization process to ensure that a fully ready sound applet exists before the sound proxy allows the rest of the Javascript implementation, including the audio indicator set, to begin interacting with the applet.

When a browser accesses and initiates the loading of an audiovisual web map application implemented using this software design, the Javascript files and the Java sound applet downloads begin asynchronously. Each of them will eventually complete but the order of those two completion events are unknown. Because the sound applet has nothing to do until the map design initiates activities for the loading of sound or MIDI files and the configuration of parameters related to the use of those sounds, it simply waits for the sound proxy to tell it what to do. As part of initializing the entire audiovisual web map, the application-specific Javascript initialization routines must tell the sound proxy where the sound applet is being loaded in the DOM and this is done by calling the sound proxy's

`ConfigureSound(`*nameInDOM*`, ... ,` *callbackFn*`)`

function.[4] The *nameInDOM* parameter identifies where the sound applet is in the DOM so that the sound proxy can start to look for the applet to determine when it is ready to handle requests. Once the sound proxy has determined that the sound applet is ready, by first determining that it can find the applet and then eventually getting a positive response to periodic polls of the applet's `isActive()`

[3]Assuming that the file defining the MIDI composition is organized as a Type 1 Stored MIDI File (SMF, see Heckroth 1996), then managing the MIDI playback at the level of tracks and channels *can* correspond to managing each instrument separately, although there are other ways to organize these files so a map author must understand the organization of a MIDI composition to be used.

[4]`ConfigureSound()` has many parameters which are of relevance only when used as part of *Nunaliit* version 1.0 (Hayes 2006). For this discussion, only those shown are relevant. All others can be left *null*.

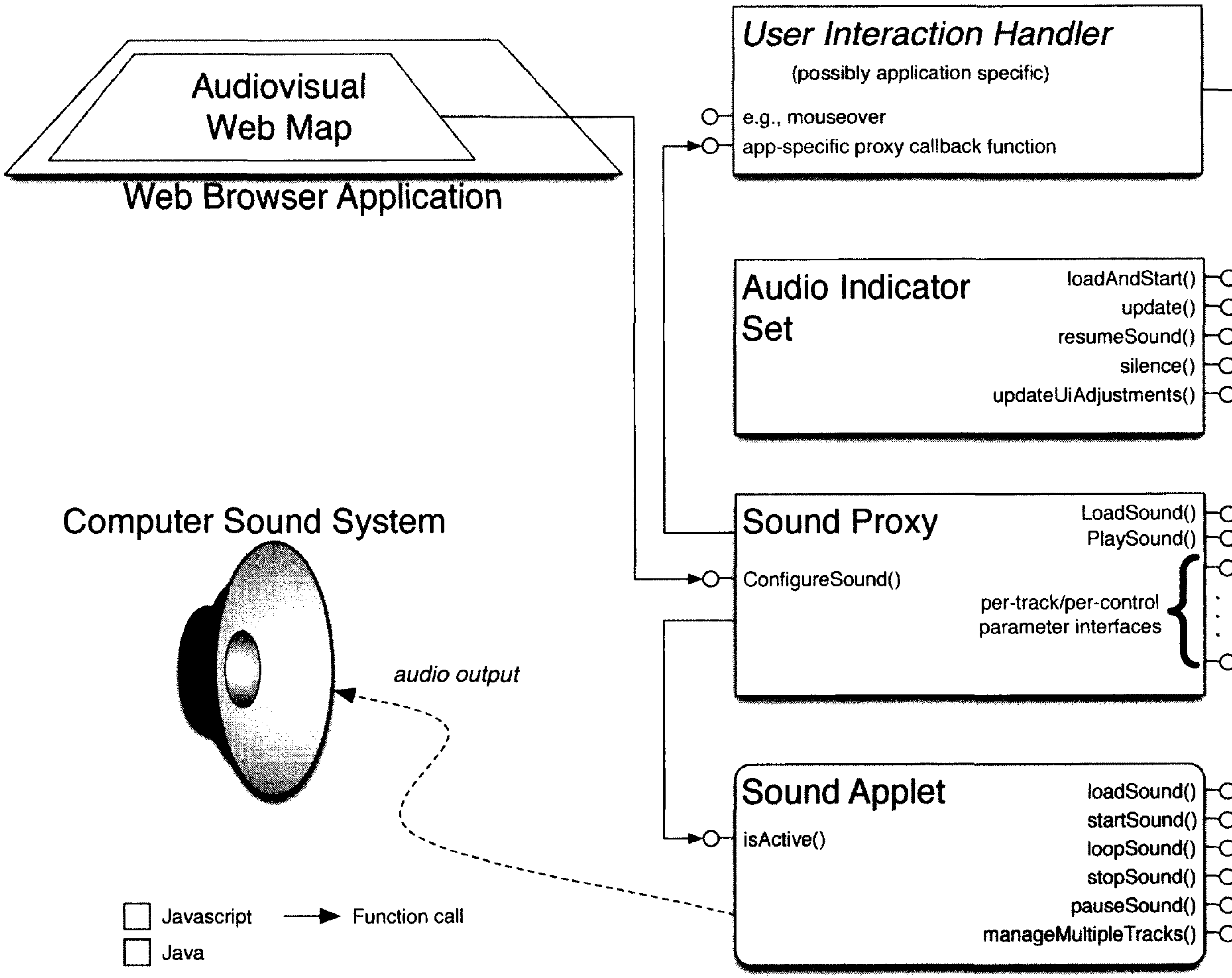


Figure F.2: Major sound subsystem software components and exemplar functional interfaces used during initialization (figure concept and design: Glenn Brauen).

function, the sound proxy calls the application-specific callback function identified as *callbackFn* in the earlier call to its `ConfigureSound()` interface. By calling the identified callback function, the sound proxy signals that it and the sound applet are both now ready for the initial configuration of the application sound design.

In addition to configuring the sound applet and proxy, the audiovisual web map application must create the audio indicator set configuration to be used once the applet and proxy are ready (see Figures 7.2 and 7.4). The audio indicator set contains a set of `midiInidicator` and `sampledAudioIndicator` configurations, each of which are associated with a *key*, a symbolic identifier encoded as a character string denoting the monitored variable associated with the configured indicator. The indicator set manages indicators within keyed arrays so that each unique key may identify a number of indicator configurations allowing for multiple sources and effect processor chains to be treated as if they are a single complex composite audio indicator. An audio indicator set is created using the function call

$$aiSet = \texttt{audioIndicatorSet}(midiSmfUrl)$$

where *midiSmfUrl* is a Universal Resource Locator (URL) identifying a MIDI stored MIDI file (SMF) containing the instrumental performances to be used for any audio indicators using MIDI instruments as their sound sources. If the audio indicators for this set all use sampled audio, then *midiSmfUrl* can be omitted. The call to `audioIndicatorSet()` returns an empty audio indicator set, identified above as *aiSet*, ready for each audio indicator to be added to it.

Each audio indicator is added to a set using the function

$$aiSet.\texttt{addIndicator}(key,\ configuredIndicator)$$

where *key* is as described above and *configuredIndicator* is either a previously configured MIDI or sampled audio indicator, to be discussed below. Subsequent calls to update audio control parameters using the audio indicator set's `update()` function, discussed above, pass in a key which is matched against the set of keys previously configured in the audio indicator set to determine which audio indicators are to be updated by that `update()` call. Multiple calls to `update()`, one for each key previously used in adding audio indicators to the set and thereby identifying a unique monitored variable, are thus needed to fully update all indicators in the set. The same *key* value can be used in multiple calls to `addIndicator()` thereby creating composite indicators containing more than one audio source, each with its own set of effect processors. For example, see Figure 7.9 in which the key 'balance' is used in two calls to `addIndicator()`, one adding the nylon string guitar and the other adding the cello to create a composite indicator for the balance of trade variable used in that application.

The audio indicators added to an indicator set are of two main types: MIDI and sampled audio. In each case, an audio indicator is created, identifying its sound source either as a track and channel of the MIDI composition previously passed to the audio indicator set as *midiSmfUrl* or as an URL identifying a sampled audio source passed directly to the audio indicator when it is created. The functions to create audio indicators are:

```
m = midiIndicator({
        trackId : track#,
        channelId : channel#
    });
s = sampledAudioIndicator({
        sourceUrl : filePathAndName
    });
```

where *track#* and *channel#* identify the MIDI channel and track, respectively, that contains the source audio definition for a MIDI indicator and *filePathAndName* identifies a sampled audio file containing the source sound for a sampled audio indicator. Each of these functions returns an audio indicator object to which audio effect processing components are added, as shown below for the sampled audio indicator *s*, using the function

```
s.addProcessingComponent(comp1, ... , compN)
```

where each of *comp1* through *compN* specify the configuration of a single audio effect processing component to be included as part of the effect processing chain for the indicator.

The audio effect processing components can be configured flexibly to control a single audio parameter based on changes to a monitored variable and are specific to each control parameter (e.g., gain), to whether a classified or unclassified representation of a variable is being used, and to the audio source type. For example, as shown in Figure 7.9, the following example would add a classifying MIDI gain control to the MIDI indicator *m*, configuring it to use the specified gain settings and to use the identified classification object:

```
m.addProcessingComponent(
    midiChannelClassifiedGain({
        classification : classification,
        gains : gainsArray
    })
);
```

Once an audio indicator has been created and the required audio processing components have been added to it, the indicator is added to the indicator set as described above. For details concerning audio processing components, see the software listed in Appendix G.

Having configured an audio indicator set, the application-specific Javascript for the audiovisual web map may then call the function

```
aiSet.loadAndStart()
```

to begin the download and initialization processes of the application's sound design but only after the sound proxy and applet are ready as indicated by the invocation of the application function identified as *callbackFn* in the call to the `ConfigureSound()` function. Once notified by the invocation of that callback function, the application may also access any sound proxy functions directly, as discussed above. As part of initializing the sound design, many of the functional interfaces of the audio indicator set, the sound proxy, and the sound applet will be used as described above in the discussion of audiovisual map sound update handling.

F.2.2 Sound Parameter Controls Implemented

Table F.1 shows the types of sound parameter controls that have been implemented in the latest version of the sound subsystem, separated according to the types of classification scheme and sound source in use. The available set of parameter controls is related more to what audiovisual map experiments were conducted and what controls were felt necessary for those than to what is possible. In particular, the fact that a specific type of parameter control was or was not included for one type of classification scheme or for one of the audio source types does not indicate that it could not be included but only that it has yet to be included. For example, pan controls for MIDI sources would certainly be possible but have not been needed in the work done to date.

Table F.1: Sampled audio and MIDI parameter controllers currently supported by the sound subsystem.

	Audio Source Type	
Classifier Type	**Sampled Audio**	**MIDI Instrument**
classifying	mute	mute
	gain	gain
	pan	–
	playback rate conversion	–
	–	pitch transposition
	play state controller	–
linearly interpolated	gain	gain
	pan	–
	playback rate conversion	–
	–	pitch transposition

By contrast, the *playback rate conversion* controller, only implemented for sampled audio sources, was designed to directly manipulate streams of sampled audio values.[5] It alters the playback of a sampled audio stream by both changing the speed at which the audio is played back, speeding it up or slowing it down, and it adjusts the pitch of the reproduced audio up or down, depending on whether the playback is sped up or slowed down. Similar effects for each of these changes are possible with MIDI sources (e.g., by combining pitch transposition and tempo[6] controllers) but the implementation is different.

The sampled audio *play state controller* is used to start and stop playback of sampled audio files in response to variable updates. The qualitative difference between a sound being played or not makes this seem most obviously applicable for use with a classified variable and thus has only been implemented for this case. To date this type of controller has been used only with short sampled audio clips intended to signify a certain condition (e.g., one or two word narrations declaring a variable state). The sounds automatically reset to play from the beginning upon repeated play commands but are usually configured so that the sound is only replayed when the variable value is changed, thereby reducing unnecessary repetition. Because playback of a stored MIDI composition maintains synchronization across tracks, the application of *play* and *stop* commands on each track has not been practical for the audiovisual map developments done to date and it seems that this would only be practical if using a store MIDI composition comprised of a single track. In such a case, there would be no need for synchronization and the same result could most probably be achieved using sampled audio.

[5]Sampled audio playback rate conversion is performed by determining at which temporal position within an input stream of samples the next desired output sample should occur, based on a control parameter adjusting the play rate as a percentage of the input sample stream's characteristic sample rate (values below 100% speeding up playback and values above 100% slowing down playback). Once the desired output position is known, the time lag between that output sample's position and those of its temporally closest neighbouring input samples are computed and those lags are used in an inverse (temporal) distance weighted interpolation of those nearest input samples to approximate the output sample value (the waveform amplitude).

[6]Tempo controls for MIDI also apply to all tracks in a stored MIDI file (SMF) so this control may be problematic to adjust, unless you are using only a single track MIDI file, alone or in conjunction with sampled audio indicators.

Appendix G: Media and Source Code CD-ROM Contents

This thesis includes a CD-ROM containing: videos showing the use and behaviour of the audiovisual map examples discussed in this thesis; and collections of source code files implementing the audiovisual mapping sound subsystem, its Application Programming Interface (API), and the main source files implementing *Canada-USA Commodity Trade 1976–2000* (Chapter 8) as an example of the use of the sound subsystem.

The CD-ROM contains videos for the map examples included in the user evaluation sessions discussed in this thesis (Chapter 6 and Appendix B) and *Canada-USA Commodity Trade 1976–2000* (Chapter 8 and Appendix H). Although the descriptions of the map examples in this thesis include Universal Resource Locators (URL) that will allow a reader to access each map on the World Wide Web, this repository of videos will allow a reader to see (and hear) the application being used even if they do not have access to a supported web browser. Current versions of the online map examples contain usage notes including statements concerning those browser with which the map examples are known to work. Over time, the web applications may become more difficult to access online if (when) contemporary browser technologies change and cease to work with these examples. In that case, these videos will provide another means of archiving these examples for future reference.

The source code files are the most recent version of the software at the time of writing and are included to clarify the discussion in this thesis of the sound subsystem software.[1] The folder structure and the purpose of the software in each is briefly outlined. Key files in each folder are identified where these would provide a good starting point in trying to understand the structure of the software in that folder.

The remainder of this section contains tables listing the contents of the CD-ROM. These tables are replicated on the CD-ROM in a file named `AnnotatedContents.pdf`.

[1] Readers interested in working with the software are encouraged to contact the author to obtain a more recent version, should one exist.

G.1 Map Example Video Captures

Table G.1 contains a list of the audiovisual map screen captures included on the CD-ROM along with each file's location and a description of what the video shows.[2]

For the interactive map examples, the location of the cursor is used to select values for the spatial variables represented acoustically. The video captures show the cursor so you can correlate its location with the simultaneous auditory and visual changes.

Sometimes it is also necessary to click the mouse button to cause the web application to select an option. To identify this in the videos, the screen capture software briefly displays a green circle, as shown, growing outward from the location of the cursor when clicked. This method for indicating when the mouse button has been clicked is used in all screen capture videos except for that of the version of *Ottawa Area Federal Election Sound Map: June 28, 2004* designed for use with the Adobe Scalable Vector Graphics (SVG) viewer plugin. This version of the map is effectively obsolete and I no longer have a computer on which it works. Because that version of the map was used in the evaluation sessions described in Chapter 6, I included an older video capture of that map that does not visually indicate when the mouse button is clicked. In that video, the mouse button is only clicked once at the beginning of the session to give control to the SVG viewer plugin, causing it to start the audio recordings looping.

[2]All of the videos were created using *Quicktime* and are viewable using a free *Quicktime* player. Depending on your system, you may need to download a player (http://www.apple.com/quicktime/, at the time of writing).

Table G.1: CD-ROM Contents: Map example audiovisual screen captures. Continuing on next page ...

Map	Example File	Description
Ottawa Area Federal Election Sound Map: June 28, 2004 (OAFESM)†	AudiovisualMapExamples/ OttawaAreaFederalElection2004/ FedElectionSoundMap_java.mov	Video capture of the election map version implemented using **standards-compliant Scalable Vector Graphics (SVG) and the Java-based cartographic sound subsystem described by this thesis to manage audio playback**. Shows **interactive selection of the map regions** and the resulting playback of leaders' speeches blended according to the vote percentages in each district as well as the **use of the legend to isolate each speech**.
	AudiovisualMapExamples/ OttawaAreaFederalElection2004/ FedElectionSoundMap_adobe.mov	Video capture of the election map version implemented using **proprietary sound extensions of the Adobe SVG viewer plugin**. Shows **interactive selection of the map regions** and the resulting playback of leaders' speeches blended according to the vote percentages in each district as well as the **use of the legend to isolate each speech**. This version of the map is **effectively obsolete** because support of that viewer plugin has been discontinued but it **is the version used during the user evaluations** described by this thesis (see Chapter 6).
Canada's Trade with World Regions: 1976–2000 (CTWR)†	AudiovisualMapExamples/ CanadasTradeWithWorldRegions/ Trade_defaultView_allSectors.mov	Video capture of the world trade map showing **interactive selection of regions and the resulting updates of the auditory representation** of spatial variables. Also shows the interactive selection of regions, **drawing of the trade bar charts**, and the **sonification of the selected bar charts** to compare, in this case, trade across regions during the same time period.

Table G.1: CD-ROM Contents: Map example audiovisual screen captures (continuted)

Map	Example File	Description
	AudiovisualMapExamples/ CanadasTradeWithWorldRegions/ Trade_defaultView_ compareTradeSectors_Asia.mov	Video capture of the world trade map **showing a comparison, visual and auditory, of the three main trade sector groupings** for Asia during the same period of time. Auditory representation clearly distinguishes Canadian trade deficit in manufactured goods from the trade surpluses Canada has with Asia during this period in food and agriculture sectors, and materials and energy sectors. Note that the same approach would allow you to compare trade in a given sector category across time periods.
	AudiovisualMapExamples/ CanadasTradeWithWorldRegions/ Trade_learnAudioView_ visualMetersAsAids.mov	Video capture of the world trade map **showing the 'learn/configure audio' panel** of the application. In particular, the visual variable indicators for the acoustically represented variables are shown varying with the auditory representation.
	AudiovisualMapExamples/ CanadasTradeWithWorldRegions/ Trade_learnAudioView_ visualMetersAsControls.mov	Video capture of the world trade map again **showing the 'learn/configure audio' panel** of the application but now showing the **use of the visual variable indicators** for the acoustically represented variables **as audio controls** used to directly manipulate the sounds as a way of learning how the acoustic representation works. Also shows a **comparison of the value of exchange variable being used with and without classification**. About 1:00 into the video, the user configures the application to use a nonclassifying audio design.

Table G.1: CD-ROM Contents: Map example audiovisual screen captures (continuted)

Map	Example File	Description
Cybercartographic Atlas of Antarctica: Territorial Claims Section (ATC)[‡]	AudiovisualMapExamples/ CybercartographicAtlasOfAntarctica/ TerritorialClaims.mov	Video capture of the map section concerning contested territorial claims showing the **interactive selection, by cursor movement, of national claims to Antarctic territory**. Selected national sectors are highlighted on the map and hyperlinks naming the claiming country in the text are highlighted as well. Selection can be accomplished by positioning the cursor over either map sectors or hyperlinked text. If a selected national claim is contested, then a narration becomes audible when that national claim is selected as discussed in Section B.3.1. Video capture also **shows additional information being accessed, by clicking on a map sector**, for both disjoint national claims and contested national claims.
Cybercartographic Atlas of Antarctica: Exploration Section (AE)[‡]	AudiovisualMapExamples/ CybercartographicAtlasOfAntarctica/ Exploration.mov	Video capture of the atlas section concerning exploration of Antarctica. The **cursor initially sweeps across the array of expedition selectors (expedition chronology)** showing the map display updates that occur with the cursor movements. Then a **selection of the expedition selectors are clicked** to show some example updates to both the map display (replacement of the transforming myth of the Island of Utopia with the relief map of Antarctica) and the multimedia information panel.

Table G.1: CD-ROM Contents: Map example audiovisual screen captures (continuted)

Map	Example File	Description
Canada-USA Commodity Trade 1976–2000†	AudiovisualMapExamples/Canada-USA_Commodity_Trade_1976-2000/CanUsaTrade_IntroAndInteractiveMap.mov	Video capture of the introduction page and interactive map page of the Canada-USA commodity trade map being used. Provides an overview of **map audiovisual updates map resulting from cursor movements** over the map; use of the **dynamic audiovisual legends** as visual learning aids and as direct sound controllers; use of **trade data selection options**; **region text detail updates** resulting from clicking on map regions.
	AudiovisualMapExamples/Canada-USA_Commodity_Trade_1976-2000/CanUsaTrade_interactive_audioConfigOptions.m4v	Video capture showing and explaining the **use and effects of the audio configuration options** built into the interactive and sequenced map pages of the application. This is a training video included in the application.
	AudiovisualMapExamples/Canada-USA_Commodity_Trade_1976-2000/CanUsaTrade_interactive_SelectorsAndSectorLegend.mov	Video capture showing and explaining the use of the **trade data selection options** (trade sector, trade flow direction, and year) in relation to the visual updates for the Trade Sectors Legend and the trade flow depiction used in the map. This is a training video included in the application.
	AudiovisualMapExamples/Canada-USA_Commodity_Trade_1976-2000/CanUsaTrade_sequenced_timeOrdered.mov	Video capture showing and explaining the **configuration and use of time-ordered map sequences** (trade data for a single region shown through time). This is a training video included in the application.

Table G.1: CD-ROM Contents: Map example audiovisual screen captures (continuted)

Map	Example File	Description
	AudiovisualMapExamples/Canada-USA_Commodity_Trade_1976-2000/CanUsaTrade_sequenced_OrderedRegions.mov	Video capture showing and explaining the **configuration and use of ordered regions map sequences** (comparison of concurrent trade data observations for a group of regions, ordered according to a selected trade variable). This is a training video included in the application.
	AudiovisualMapExamples/Canada-USA_Commodity_Trade_1976-2000/CanUsaTrade_sequenced_importExport.mov	Video capture showing and explaining the **configuration and use of import/export map sequences** (comparison of imports and exports with other data selections kept constant: region, year, and trade sector). This is a training video included in the application.

† Design and Production by Glenn Brauen.

‡ Courtesy of Geomatics and Cartographic Research Centre, Carleton University (Caquard et al. 2008; Pulsifer et al. 2007, 2005).

G.2 Source Code

Table G.2 lists some of the key folders and source code files included on the CD-ROM accompanying this thesis. There are three main folders included: *SoundApplet*, *JavaScript*, and *canusa*.

The folder src/SoundApplet contains the implementation of the sound applet that manages audio for the cartographic sound subsystem described in this thesis (see Chapter 7 and Appendix F). The folder src/JavaScript contains key files implementing the sound applet application programming interface (API), the variable monitoring, and the audio indicator design that uses that API. The folder src/canusa contains the implementation of the audiovisual map example, *Canada-USA Commodity Trade 1976–2000*, discussed in this thesis as an example of the use of the sound subsystem (see Chapter 8 and Appendix H).

The inclusion of these files is intended to be illustrative and there is no intent that the software as included is a complete version of the components that could be installed and run. To accomplish that, external libraries, the database of Canada-USA trade values, and the software for the application server that interprets application database query requests are also required. The main external dependencies required to build the SoundApplet are described in the file src/SoundApplet/readme_external_libs.txt included on the CD-ROM. The main file of the implementation of *Canada-USA Commodity Trade 1976–2000*, src/canusa/index.html, includes a full list of all the code dependencies required to run the example. The database of Canada-US trade flows was produced by Statistics Canada and licensed to Carleton University's Geomatics and Cartographic Research Centre for research purposes and is not redistributable under that license. Readers interested in obtaining full, possibly updated versions of the open source software components are encouraged to contact the author.

Tables G.3 and G.4 provide listings of the sources files included on the CD-ROM along with line counts for each. Table G.3 includes all of the Java source files for the Sound Applet implementation and Table G.4 includes javascript files for the sound subsystem API and parts of the *Canada-USA Commodity Trade 1976–2000* implementation, as described in more detail in Table G.2.

Table G.2: CD-ROM Contents: Brief description of key source code folders and files. Continuing on next page ...

Main Folder	Key Files	Description
src/ SoundApplet	This folder contains the Java implementation of the SoundApplet that manages the real-time audio processing on behalf of the cartographic sound subsystem described by this thesis (see Chapter 7 and Appendix F). This was used as the sound subsystem in the *Nunaliit Cybercartographic Atlas Framework, Version 1,0* and was a collaborative effort by the author of this thesis and J. P. Fiset. For some of the files below, key contributions by J. P. Fiset are noted. Because each of the files listed for the sound applet are included in the ca.carleton.gcrc.atlas Java package, each of the files discussed below is actually in the folder src/SoundApplet/ca/carleton/gcrc/atlas/.	
	SoundApplet.java	Implements the main sound applet interfaces for loading and unloading audio files, starting and stopping playback of files, adjusting playback parameters such as gain, pan, etc. Provides capabilities for asynchronous event notifications to be registered against audio files (e.g., END notifications).
	TrackManager.java	Implements sampled audio track management, including allocation of tracks, managing audio parameter control adjustment.
	MidiManager.java	Implements MIDI audio management, including loading and unloading of MIDI compositions, and playback control adjustment.
	tracks [folder]	Contains the implementation of sampled audio tracks, including allocation and management of (input) audio streams, format conversions, management of asynchronous event detection (end of file), and management of requests for available (output) data lines.

Table G.2: CD-ROM Contents: Brief description of key source code folders and files (continued)

Main Folder	Key Files	Description
	tracks/versatile/VersatileTrack.java	Implements sample audio stream suitable for use as a looping audio stream (no gap restart) and as an audio clip (non-streamed). Allocates (input) audio streams, required audio stream resources for effects processing, and pooled (output) data lines as required. Based on earlier, separate StreamTrack and ClipTrack implementations, J. P. Fiset designed this class initially to handle allocation of pooled data lines (obsoleting the earlier StreamTrack class). Glenn Brauen subsequently updated this class to also replace and obsolete the earlier ClipTrack class).
	audiostream [folder]	Contains the implementation of input sampled audio streams including the capabilities to send audio data to a player, restart a stream, play a stream in a loop, detect and report end of file, and close and input stream.
	audioStreamDSP [folder]	Contains the implementation of audio effects processing using a chained model of effects processors. A buffer containing audio data is passed from processor to processor down the chain with the altered audio then passed to the (output) data line.
	dataline [folder]	Contains the implementation of the output data line, managed within a pool of data lines for flexible assignment of audio streams to data lines as needed. J. P. Fiset designed the pooled data line management.
src/ JavaScript	This folder contains Javascript software components designed to be reused across audiovisual web map applications. They are not specific to any particular map. This folder includes implementations of software components used to implement the audio designs described in this thesis, key files of which are identified below.	

Table G.2: CD-ROM Contents: Brief description of key source code folders and files (continued)

Main Folder	**Key Files**	**Description**
	widgets/audioRepresentation/ AudioIndicator.js	Implements the audio indicator sets described in this thesis (see Chapter 7 and Appendix F).
	widgets/audioRepresentation/ AudioProcessingComponents.js	Implements the *midiIndicator* and *sampledAudioIndicator* designs described by this thesis (see Chapter 7 and Appendix F) and specific audio parameter controls (e.g., gain, pan, sample rate conversion controls) used by each of these.
	widgets/audioRepresentation/ soundProxy.js	Implements the sound proxy described by this thesis (see Chapter 7 and Appendix F). J. P. Fiset did the original implementation based on his idea to batch parameter updates from the browser to the sound applet for better performance. Glenn Brauen later added MIDI support, more parameter controls, support for effect processing chains, and asynchronous event notifications from the applet to the browser.
	widgets/mapLogic/classification.js	Implements both the classifying and linear interpolation methods of classification used with monitored variables to update the audio parameters control by the audio indicator designs described in this thesis (see Chapter 7 and Appendix F).
	util/VariableMonitor.js	Implements the variable monitor design used to update the audio parameters controlled by audio indicators described in this thesis (see Chapter 7 and Appendix F).
src/canusa	This folder contains the implementation of *Canada-USA Commodity Trade 1976–2000*. HTML, JavaScript, and media files are all included.	
	index.html	Main file. This contains the initial page layout (although much is loaded dynamically) and the full list of included source files implementing the behaviour of the application.

Table G.2: CD-ROM Contents: Brief description of key source code folders and files (continued)

Main Folder	Key Files	Description
	CANUSA_main.js	Initialization of the application begins here.
	CANUSA_AudioIndicators.js	Implements the creation, initialization, loading, and updating of the auditory indicators used by this application.
	CANUSA_AudioMeters.js	Implements the creation, initialization, and management of the dynamic audiovisual legends used by this application.
	CANUSA_CartographicSequencer.js	Implements the cartographic sequencing used by the sequenced maps page. See also JavaScript/canusa/topicContent/animated.js which implements the details for each type of map sequence supported, including the computation of the comparative totals used for the relative auditory measures (e.g., the Canada-USA trade totals needed for determining each region's contribution percentage).
	CANUSA_FormUserInterface.js	Implements dynamic creation of forms used to select trade options, sequence types for animation, and containing the dynamic audiovisual legends and audio configuration panels in the application interface.
	CANUSA_MapUserInteractions_common.js	Implements common elements of the map user interaction handling. See JavaScript/canusa/topicContent folder for details of customization done for interactive vs. Sequenced map interactions.
	CANUSA_SelectableTradeDataOptions.js	Implements a data holder class, specifying the details of one trade data observation (year, sector, focus region, etc.). An array of these can then be the basis of a map sequence.

Table G.2: CD-ROM Contents: Brief description of key source code folders and files (continued)

Main Folder	Key Files	Description
	CANUSA_SqlDataRequests.js	Implements the database interactions required to get trade data from the database (using asynchronous JSON requests) and store it in the application tables.
	CANUSA_TextDetails.js	Implements the creation of the application text panel (when a user clicks on a map region).
	CANUSA_TopicNavigation.js	Implements the tabbed navigation between the overview, interactive, and sequenced application pages. See topicContent folder for details of each loaded page.
	CANUSA_TradeData.js	Implements trade data storage and computation for the application.
	CANUSA_UserInterfaceOptions.js	Implements a data holder class, specifying the details of user options not related to the selection of gain data (e.g., sequencing order, sequencing delay).
	CANUSA_VisualTradeFlowLayer.js	Implements the visual thematic map layer, including display and update of legends, and the display and animation of trade flow lines and pie charts (see also JavaScript/widgets/functionalAnimator.js and JavaScript/widgets/SvgProportionalSizePieChart.js).

Table G.3: Java Source Files for Sound Applet, with line counts. Files are in the folder src/SoundApplet/ca/carleton/gcrc/atlas/ on the CD-ROM.

Lines	File name
49	audiostream/AudioFiniteBufferedStream.java
5	audiostream/AudioPlayer.java
374	audiostream/AudioReloadingStream.java
55	audiostream/AudioStream.java
101	audiostream/SupportingFiniteBufferedStream.java
54	audiostream/SupportingStream.java
73	audioStreamDSP/AudioStreamDSP.java
248	audioStreamDSP/AudioStreamDSPConfiguration.java
46	audioStreamDSP/AudioStreamDSPConfigurationException.java
200	audioStreamDSP/ByteBufferDSPChain.java
40	audioStreamDSP/DSPChainProcessingParameters.java
268	audioStreamDSP/FloatSampleBufferRateConverter.java
106	dataline/DataLinePool.java
56	dataline/DataLineProxy.java
450	dataline/DataLineProxyImpl.java
5	dataline/DataLineRequestor.java
88	FadeTimerParent.java
365	FadeTimerTask.java
84	logger/Logger.java
343	MidiManager.java
371	OnePassMidiTransformer.java
84	OnePassMidiTransformerReceiver.java
76	OnePassMidiTransformerTransmitter.java
904	SoundApplet.java
53	SoundException.java
142	SoundTest.java
603	StandardMidiFilePlayer.java
1182	TrackManager.java
40	TrackManagerTrackType.java
83	tracks/AudioTrackAsyncDetector.java
37	tracks/AudioTrackAsyncEventType.java
47	tracks/AudioTrackAsyncRecipient.java
273	tracks/AudioTrackBase.java
5	tracks/AudioUtil.java
155	tracks/conversion/AudioStreamConverter.java
25	tracks/conversion/InternalConverter.java
54	tracks/conversion/Mp3Converter.java
56	tracks/conversion/OggVorbisConverter.java
49	tracks/conversion/SystemConverter.java
118	tracks/IAudioTrack.java
7	tracks/Playable.java
62	tracks/PlayThread.java
11	tracks/ResourceRequest.java
54	tracks/ResourceRequestSimple.java
169	tracks/versatile/VersatileTrack.java
7670	**Total**

Table G.4: Javascript Source Files for Sound Subsystem API and *Canada-USA Commodity Trade 1976–2000* implementation, with line counts. Files are in the folder src on the CD-ROM.

Lines	File name
638	canusa/CANUSA_AudioIndicators.js
444	canusa/CANUSA_AudioMeters.js
390	canusa/CANUSA_CartographicSequencer.js
804	canusa/CANUSA_FormUserInterface.js
190	canusa/CANUSA_main.js
361	canusa/CANUSA_MapUserInteractions_common.js
112	canusa/CANUSA_SelectableTradeDataOptions.js
1221	canusa/CANUSA_SqlDataRequests.js
152	canusa/CANUSA_TextDetails.js
424	canusa/CANUSA_TopicNavigation.js
319	canusa/CANUSA_TradeData.js
84	canusa/CANUSA_UserInterfaceOptions.js
569	canusa/CANUSA_VisualTradeFlowLayer.js
104	canusa/data/MaximumFlowValue.js
95	canusa/data/TradeFlowObservationTable.js
892	canusa/topicContent/animated.js
176	canusa/topicContent/interactive.js
20	canusa/topicContent/overview.js
17	JavaScript/util/common.js
132	JavaScript/util/number_format_utilities.js
103	JavaScript/util/VariableMonitor.js
72	JavaScript/widgets/audioRepresentation/AudioIndicator_listenerBindings.js
62	JavaScript/widgets/audioRepresentation/AudioIndicator_uiAdjustments.js
629	JavaScript/widgets/audioRepresentation/AudioIndicatorSet.js
1580	JavaScript/widgets/audioRepresentation/AudioProcessingComponent.js
1456	JavaScript/widgets/audioRepresentation/soundProxy.js
159	JavaScript/widgets/functionalAnimator.js
275	JavaScript/widgets/mapLogic/classification.js
706	JavaScript/widgets/SvgAudioEnabledMeter.js
256	JavaScript/widgets/SvgContext.js
271	JavaScript/widgets/SvgFlowLine.js
308	JavaScript/widgets/SvgProportionalSizePieChart.js
234	JavaScript/widgets/SvgPSPieChartLegend.js
13255	**Total**

Appendix H: Canada-USA Commodity Trade 1976–2000: Design Details

This appendix provides complete definitions of the variables represented acoustically in *Canada-USA Commodity Trade 1976–2000* and provides design overviews for each of the auditory indicators used in the application. This appendix expands on the description of the application provided in Chapter 8.

H.1 Definition of Variables Represented Acoustically

Three variables are represented acoustically as discussed in Chapter 8. Two of these are updated as a result of the selection of the current region and associated trade selection options (sector, year) based on cursor movements in the interactive map and based on preconfigured sequence options in the sequenced maps page. These two variables, as discussed in more detail below, reflect the currently selected region's value of Canada-US trade and the region's balance of trade in Canada-US trade. The third variable, used only in the sequenced maps page and particularly intended for use with the time-ordered sequence type, varies only with the year defined for the currently represented trade flows and classifies the current value of 1$ Canadian as adjusted for inflation.

The definitions of each of these variables are given below. The design of the audio indicators used to represent the current value of each of these variables is discussed in Section H.3.

Region's value of Canada-US Trade, as a ratio of total: The value of the currently selected region's Canada-US trade as a ratio compared to the value of total Canada-US trade, computed according to the current sector and trade flow (import or export) selections for the selected region.

$$ratio = \frac{\text{Region's Canada-US Trade Value}}{\text{Total Canada-US Trade Value}} \tag{H.1}$$

Canadian:US Exports Ratio for Region: A computed ratio of imports compared to exports for the selected region that, depending on which of import or export values are greater, gives the lesser flow as a scaled ratio offset from 0 in the range

$-1.0 \ldots 1.0$, showing by how much imports are less than exports (positive) or exports are less than imports (negative). If a Canadian region is selected, then the ratio is defined by the following function.

$$ratio = \begin{cases} \text{if exports} < \text{imports} : -1.0 + \left(\dfrac{\text{exports}}{\text{imports}}\right) \\ \text{if imports} \leq \text{exports} : 1.0 - \left(\dfrac{\text{imports}}{\text{exports}}\right) \end{cases} \tag{H.2}$$

Visually, it may help to look at the *balance of trade* dynamic audiovisual legend (Figure 8.2a) and to realize that the scale of that legend matches this function. The values shown on that legend range from -1.0 (bottom, implies no Canadian exports) to 1.0 (top, implies no exports from the United States) and the base line for the legend is at 0 causing the bar chart to show larger values when a region's trade is more out of balance.

If a division of the United States is selected, Equation (H.2) is used with reversed polarity (negatives become positives and vice versa) to match the fixed polarity of the *balance of trade* dynamic audiovisual legend. Again, compare the look of the *balance of trade* legend with the above function to understand this.

Consumer Price Index Value of $1 (1976 = 1.00): Used if the *adjust for inflation* option is selected in the configuration of sequenced maps (although probably only really useful for time-ordered sequences). Represents the value of a dollar during the currently selected year in 1976 dollars, using Statistics Canada's Consumer Price Index (CPI).

H.2 Trade Data Selection Options

In January 2009, requirements for the *Canada-USA Commodity Trade 1976–2000* application had been decided:

- Support for both interactive and sequenced maps, although I was still using the term 'animated maps' at this time.
- Coordinated visual and auditory representation of trade: absolute trade value shown visually with complementary relative measures of the currently selected region represented acoustically.
- Variables to be represented acoustically: percentage of Canada-US trade contributed by the selected region; an indication of the region's trade balance; and for time-ordered sequences an indication of the value of a dollar adjusted for inflation, if the user selected the *adjust for inflation* option. The variable definitions shown above were not yet finalized and, in particular, the definition of the trade balance indication for the currently selected region changed over time.

- Ability for users to select trade data for any year in the range 1976–2000, aggregating all trade sectors or refining the selection to show only one of the top-level trade sectors (Table 8.1), and selecting either imports or exports (Figure H.1b).
- Ability for users to select either a Canadian region or a division of the United States as the current region using the interactive map. Region and division definitions are given in Tables H.1 and H.2.
- Ability for users working with the interactive map page to see summary tables of trade values for a mapped region by clicking on it on the map, based on the current selections of trade sector, year, and whether or not *adjust for inflation* is selected (Figure H.1b).
- Ability for users working with the animated maps page to configure and run three types of animated sequences:
  - **Time-ordered**: a sequence of maps for a single region with fixed trade sector and trade flow direction options (Figure H.1c).
  - **Ordered regions**: a comparison of trade flows across regions (choose to compare all regions of Canada or the United States) with fixed time, sector and trade flow directions options and ordered (for presentation time) by ascending value of the regions' Canada-US trade or trade balance indicator value (Figure H.1d).
  - **Imports/exports**: A comparison of imports and exports for a single region with fixed time and trade sector options (Figure H.1e).
- Scalability of the application by partitioning data requirements in the client (web browser application) based on use of a database and a query facility that allowed only the data required to show all region's trade flow for the current year and trade flow direction (imports or exports) on the interactive map to be loaded. Similarly, the animated map should only load the data required to show the sequence configured according to user selections.

With this set of requirements as the basis for design, I programmed the application in Javascript using Scalable Vector Graphics (W3C 2010) to draw the visual map and dynamic audiovisual legends, HyperText Markup Language and Cascading Style Sheets to create the web page layout and configuration forms, and the cartographic sound subsystem described in Chapter 7 and Appendix F for the sound design.

Figure H.1a shows the layout of the *Canada-USA Commodity Trade 1976–2000* map pages (interactive map and sequenced maps pages), which are divided into three main parts: map; data configuration and display panel (top right, with detailed views shown for each of the interactive map and sequenced maps options in Figures H.1b–e); and audio configuration and dynamic audiovisual legends (bottom right). In the interactive map page, the configuration and display panel allows a user to select the trade options they want to examine (year, sector, flow direction)

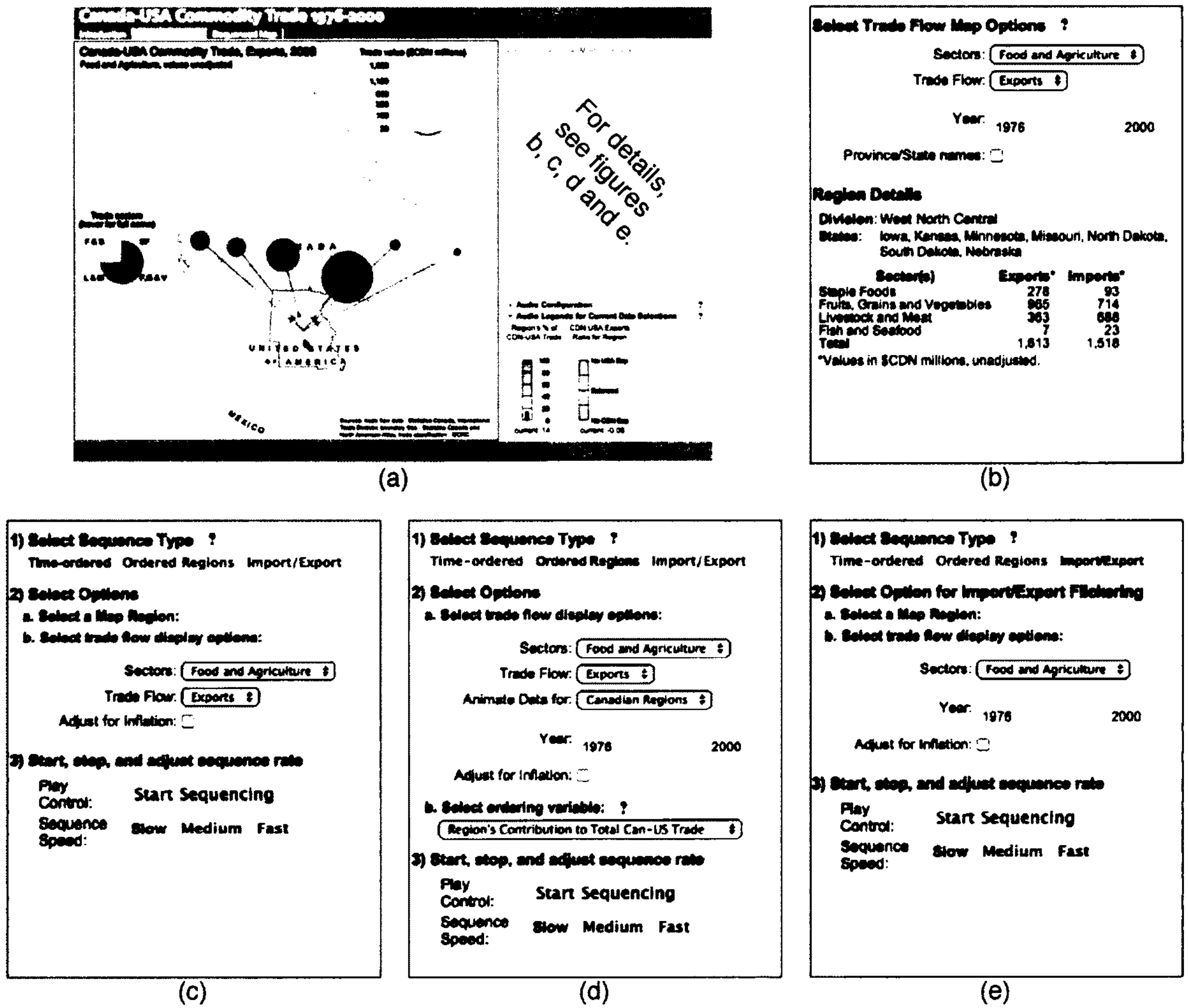


Figure H.1: Canada-USA Commodity Trade 1976–2000: Application page layout (a) and trade selection and sequence configuration options for: b) interactive maps (showing options and text display of the current region's details); c) time-ordered sequences; d) ordered regions sequences; and e) import/export sequences. Selection panels appear in greyed out portion of (a). Figure design and production: Glenn Brauen.

Table H.1: Canada-USA Commodity Trade 1976–2000 Canadian region definitions. Definitions by Statistics Canada, International Trade Division, except Alberta separated from Prairie Provinces for this research.

Region	Province(s)
British Columbia	British Columbia
Alberta	Alberta
Prairie Provinces	Manitoba, Saskatchewan
Ontario	Ontario
Quebec	Quebec
Atlantic Provinces	New Brunswick, Newfoundland and Labrador, Nova Scotia, Prince Edward Island
Territories	Northwest Territories, Nunavut, Yukon

and provides space in which region details are displayed when a user clicks on a map region (H.1b). In the sequenced maps page, the configuration and display panel enables a user to select a subset of the trade selection options, depending on the type of sequence selected, and to configure options that control how the sequence runs (e.g., ordering controls for the ordered regions type of sequence).

The audio configuration and legends portion of the application layout (lower right) contains two main sections: one containing a set of dynamic audiovisual legends each of which is associated with one of the audio indicators used in the design (Figure 8.2a); and one containing a set of sliders to be used to adjust gain and pan for each audio indicator (Figure 8.2b).

H.3 Final (or at least provisional) Sound Design

The final (for now) designs,[1] in schematic form, for each of the three audio indicators are shown in Figures H.2, H.4, and H.6. The predominant sound in the design is provided by the indicator for the value of the currently selected region's Canada-US trade. This sound plays continuously while the map is actively in use, either because a user has positioned the cursor over a map region on the interactive map or because a user has configured and started an animation sequence running using the sequenced maps page. The other audio indicators, *balance of trade* and, when in use, *value of $1*, sound briefly to mark a change in value of the associated variable but are otherwise silent.

[1]This describes the sound designs as currently installed at http://atlas.gcrc.carleton.ca/tdb/canusa. Although I have no plans to modify these designs, 'final' may be too strong a term.

Table H.2: Canada-USA Commodity Trade 1976–2000 United States division definitions. Definitions by Statistics Canada, International Trade Division.

Division	State(s)
Pacific	Alaska, California, Hawaii, Oregon, Washington
Mountain	Arizona, Colorado, Idaho, Montana, Nevada, New Mexico, Utah, Wyoming
West North Central	Iowa, Kansas, Minnesota, Missouri, Nebraska, Nebraska, North Dakota, South Dakota
West South Central	Arkansas, Louisiana, Oklahoma, Texas
East North Central	Illinois, Indiana, Michigan, Ohio, Wisconsin
East South Central	Alabama, Kentucky, Mississippi, Tennessee
New England	Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont
Middle Atlantic	New Jersey, New York, Pennsylvania
South Atlantic	Delaware, District of Columbia, Florida, Georgia, Maryland, North Carolina, Puerto Rico, South Carolina, U.S. Virgin Islands, Virginia West, Virginia

H.3.1 Value of Trade

The audio indicator for the value of the currently selected region's Canada-US trade loops a recording composed primarily of the sound of a train at a level crossing,[2] with the recording's playback rate controlled by the variable computed using Equation (H.1), as shown in Figure H.2. The variable, containing values in the range $0 \ldots 1$, is used unclassified[3] to interpolate a playback rate, as a percentage of the original recording's sampling rate, between 25% (playback takes four times as long as originally recorded) and 400% (playback takes one quarter the

[2]The train recording was done by Robert Gracek of Non Stop Art, Tychy, Poland and posted for download on the Freesound Project (http://www.freesound.org, file name: 78388__Robinhood76__01229_slow_cargo_train_2.wav). It is licensed under a Creative Commons Sampling Plus 1.0 license (http://creativecommons.org/licenses/sampling+/1.0/). For use with the map, the recording has been aggressively processed using a multiband compressor to take out a lot of low- and mid-range frequencies, leaving the rhythm of the train without the recording overpowering lower cost computer speakers or creating an oppressive background sound for the map.

[3]A *classified* variable is converted from its input range into an integer variable representing a class index in the range $0 \ldots n\text{-}1$, where n is the number of categories in the classification scheme. The audio effect processors in the design are each preconfigured to contain arrays of values to be used for the value of the acoustic parameter being controlled by that processor, with the computed class index used to index the control parameter arrays in all effect processors that are part of the indicator. An *unclassified* variable is linearly interpolated to compute an integer percentage value in the range $0 \ldots 100$, indicating the offset of the current value from the lowest value of the input range. The audio effect processors in the design are each then updated and use the integer percent value against a configured control parameter range to compute a new value for the acoustic parameter being controlled by that processor.

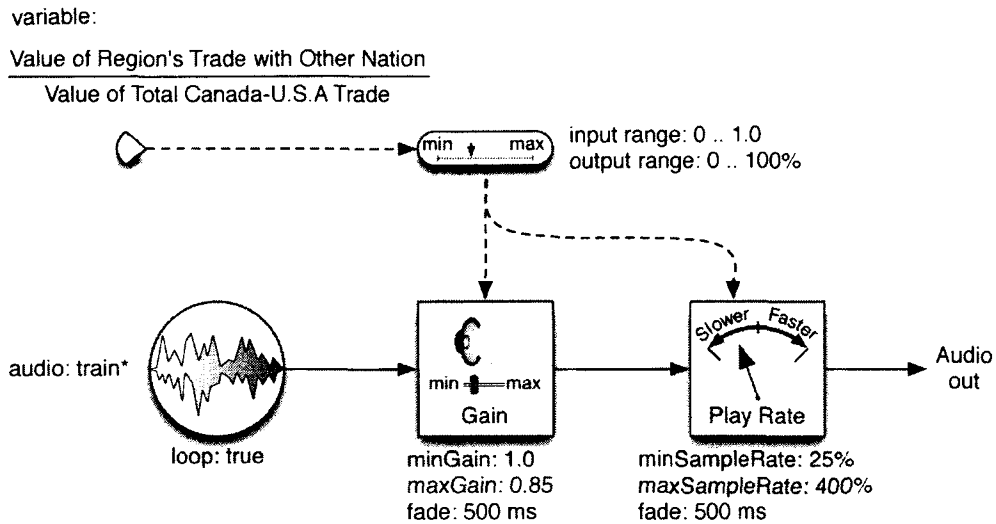


Figure H.2: Audio indicator design: Region's contribution to Canada-US trade, as an unclassified ratio (figure design and production: Glenn Brauen).

time required to originally record the sound), as shown in Figure H.3 which depicts control parameter graphs for this indicator. This emulates the sound of a train speeding up or slowing down, controlled by the ratio of the value of the current region's Canada-US trade compared to the total, creating a design based on the metaphor *'faster means more.'* As also shown in Figure H.2, a gain control is used to counteract the perception of increased volume caused by the higher frequencies produced when the recording is sped up. The gain control is configured with a control parameter range that sets the gain to 1.0 (maximum) when the input ratio is at minimum (0) and reduces the gain to 0.85 linearly across the full range of input values, slightly lowering the volume as the frequencies in the sound are increased (Figure H.3). Both the gain and playback rate controls are set to fade to a newly configured control value over a period of 500 milliseconds, causing a very noticeable period during which the train sound is slowed or sped up depending on the previous and new control values.

Because the original recording of the train contains lengthy sections of what sounds like uninterrupted white noise when played back at the lowest possible playback rate, another sound was mixed with it to create some acoustic events at these low playback rates. The current design mixes the train recording with a secondary recording, by Katarina Soukoup, of the sound made by pieces of ice in the Arctic Ocean sliding against each other.[4] The idea to use a secondary sound devel-

[4]The sound was recorded by Katarina Soukoup, a Montreal-based artist, and posted on the SoundTransit website (http://www.soundtransit.nl, file name: 0153.Katarina_Soukup.Tasiujaq_Nunavut.Ice_Squeek.mp3). It is licensed under a Creative Commons *Commons Deed Attribution 2.0* license (http://creativecommons.org/licenses/by/2.0/legalcode) meaning that it is freely available for copying, distribution, display, performance, and for creating derivative works.

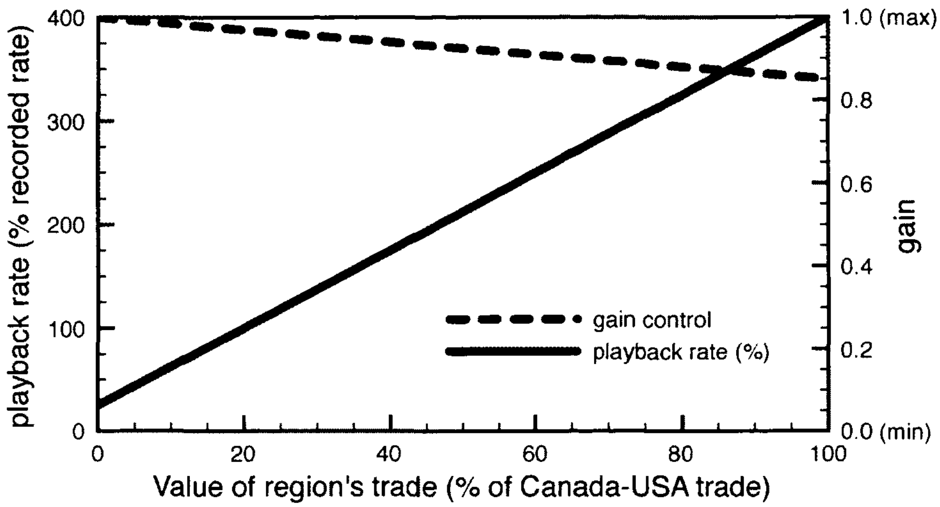


Figure H.3: Audio control parameters: Region's *value of exchange* (figure design and production: Glenn Brauen).

oped out of earlier experiments using fragments of Iannis Xenakis's 1958 composition *Concrete PH* which worked very well in creating an effective audio indicator based on playback rate adjustment because of the density of acoustic events it contains. Because the Xenakis composition, on its own, did not relate metaphorically to trade in the way I desired, I decided to mix the composition together with the train recording, which was effective across the range of playback rates and connected with the data thematically. However, because of copyright concerns with using the Xenakis composition, I eventually found the recording of the ice sheets sliding together and altered the design, replacing Xenakis's composition with the recording by Soukoup.

This combination of sounds seems to work reasonably well as a background for the map, possibly because it has a significant component of white noise, but still provides a fairly distinct feeling of rhythm which, after some time spent listening, allows reasonably small changes in the underlying variable to be detected by sound alone.

H.3.2 Trade Balance

The *balance of trade* audio indicator combines the playback of two audio clips, balancing the gain settings of the two to indicate which of Canadian exports or American exports are predominant according to user configured trade options (sector, year, and region). The current design, as shown in Figures H.4 and H.5, uses two audio recordings in a symmetric manner, a clip of a steam burst associated with

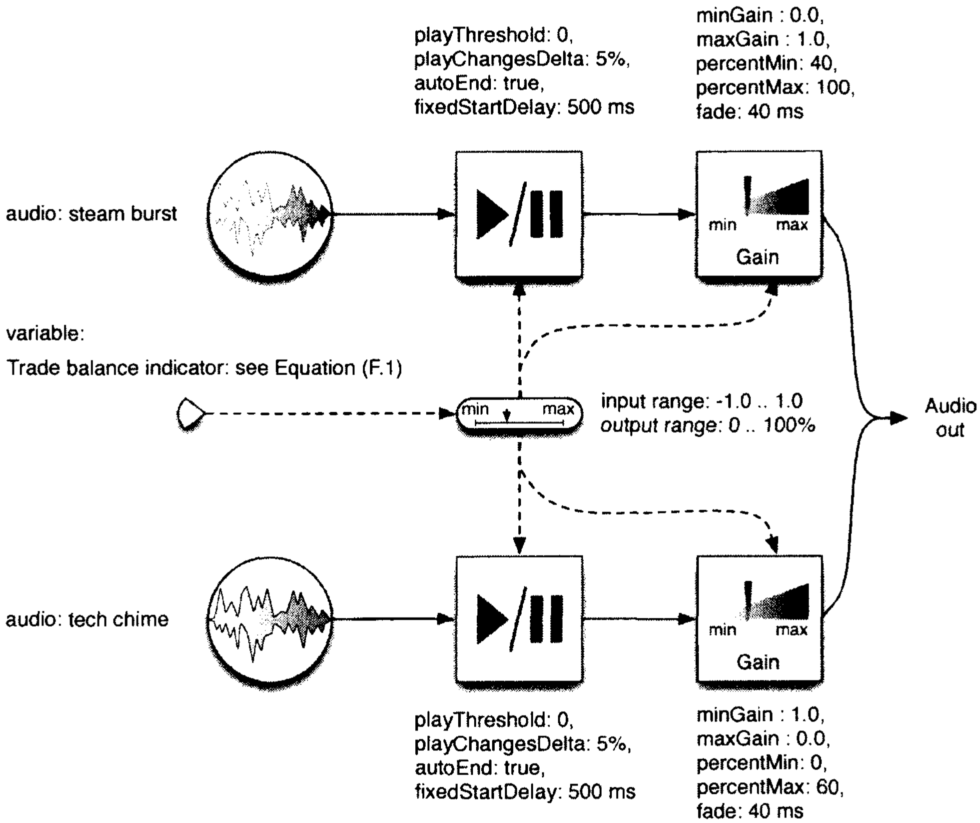


Figure H.4: Audio indicator design: Region's balance of Canada-US trade (figure design and production: Glenn Brauen).

Canadian exports to the United States and a clip of a computer start-up chime associated with US exports to Canada.[5]

The indicator design, as shown in Figure H.4, uses a pair of gain controls that linearly interpolate the gain settings for each of the two audio clips over the full gain range (0.0 ... 1.0) with symmetric cut-off gain sills configured as shown in the control graphs depicted in Figure H.5. In addition, the indicator design, as shown by the play state processing components in Figure H.4, starts each of the audio clips paying any time the underlying variable changes by at least 5% from the previous setting (`playChangesDelta` setting) but delays the clip start by 500 ms (`fixedStartDelay` setting). Ignoring variable changes of less than 5% dampens the activity of the indicator, reducing the number of times the audio clips play if a sequence of very small variable changes are detected. Delaying the audio clips'

[5]The steam burst is modified from a recording posted on freesound.org (http://freesound.org) by user pengo_au. It is licensed under a Creative Commons Sampling Plus 1.0 license (http://creativecommons.org/licenses/sampling+/1.0/). The computer start-up chime is a sample from Microsoft's Windows 95 start-up sound and is used under fair dealing provisions of Canada's copyright act.

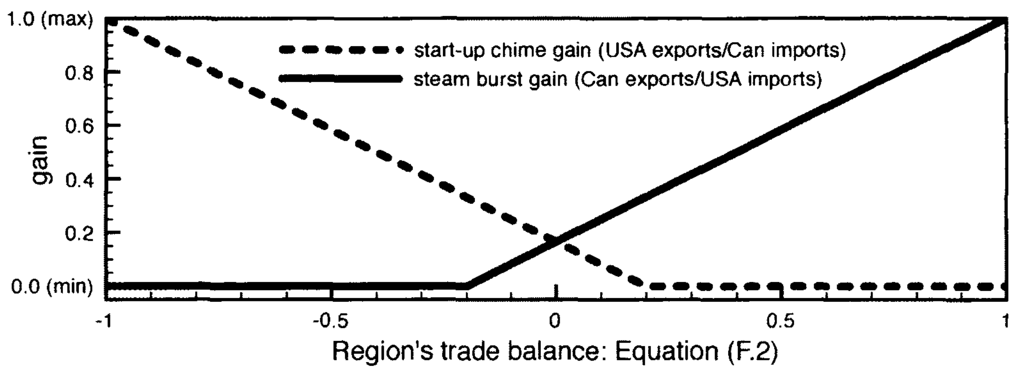


Figure H.5: Audio control parameters: Region's *balance of trade* (figure design and production: Glenn Brauen).

playback allows the gain change adjustment of the *value of exchange* audio indicator, which is also designed to fade over a period of 500 ms, to be completed before this indicator sounds, thereby staging the two indicators to work well together.

H.3.3 Inflation: Sequenced Maps (Especially Time-ordered)

Canada-USA Commodity Trade 1976–2000 allows a user to choose to display trade values that are computed using unadjusted values or that are inflation-adjusted using Statistics Canada's Consumer Price Index (CPI), with the base year set to 1976 to match the earliest trade observations in the Canada-US trade data. When inflation-adjusted data is selected, an audio indicator is enabled as part of the sequenced maps pages that classifies the inflation-adjusted value of a current dollar, designed especially to be used with time-ordered map sequences.

The *value of $1* audio indicator divides the possible range of inflation-adjusted values into classes representing $0.10 ranges, beginning with class 0 defined as $\$0.30 < x \leq \$0.40$ and continuing up to class 6 defined as $x > \$0.90$. Each class is associated with a brief vocal narration that voices the class range, as shown in Figure H.6, and is controlled by a play state audio component that starts playing the associated sound only if the underlying variable is classified into the range associated with that class.[6] When selected to be played, the class audio for the current

[6]The notation "playFlags: (class = id)" in Figure H.6 is shorthand to express the idea that the audio for a specific class is only played if the current value is in the range defining that class. The `playFlags` configuration item is, in fact, an array containing an ordered set of flags, each one corresponding to a class defined by the classification scheme and indexed by the class number. Each of the play state controls then contains an array of play flags which contains exactly one `true` value, entry n for class n, with all other entries set to false, thereby identifying the logical cases in

value of $1 begins to play after a delay of 1150 ms (`fixedStartDelay` configuration) which stages this indicator to be heard after both the *value of exchange* audio indicator has completed its transition and after the *balance of trade* audio indicator has completed playing its sounds. This classifying audio indicator retains the default behaviour of only starting to play an audio file if the underlying variable's value has changed such that the previous value and the current value fall into different classes (i.e., the change crosses a class boundary).

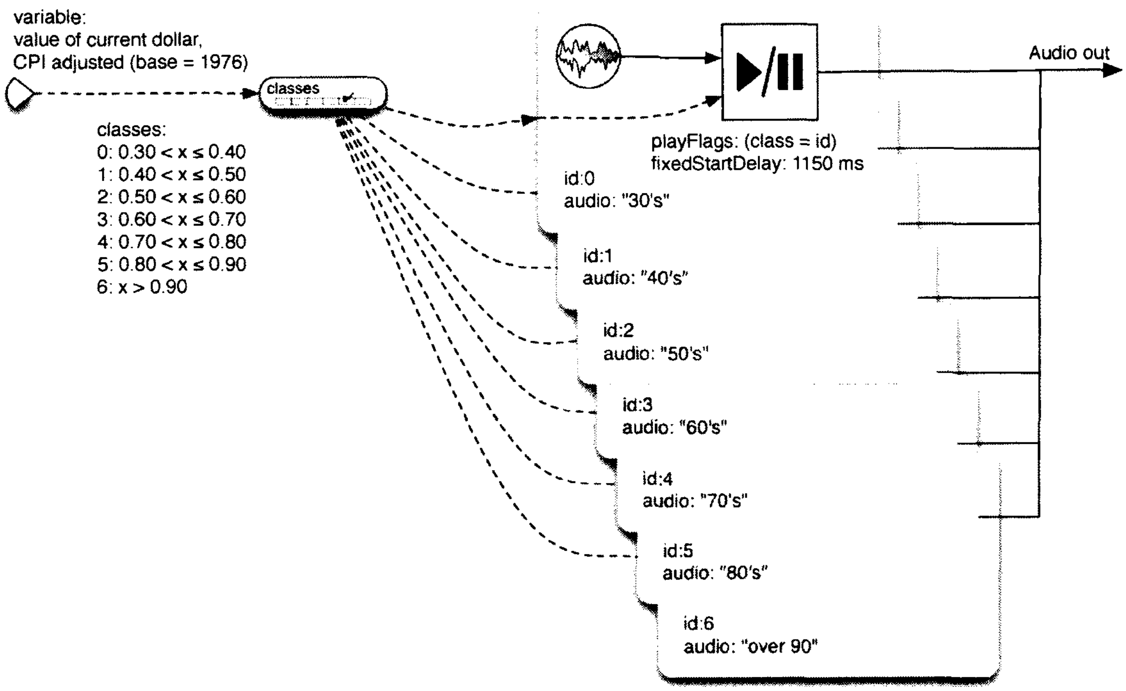


Figure H.6: Audio indicator design: Inflation adjusted value of Canadian $1, adjusted using Statistics Canada's Consumer Price Index (base=1976) (figure design and production: Glenn Brauen).

which each audio file will be played.