Listen Reader:
An Electronically Augmented Paper-Based Book

Maribeth Back, Jonathan Cohen, Rich Gold, Steve Harrison, Scott Minneman

Xerox PARC
3333 Coyote Hill Road
Palo Alto, CA USA
+1 650 812 4726
back, harrison, minneman, richgold@parc.xerox.com; cohen@dnai.com

ABSTRACT

While predictions abound that electronic books will supplant traditional paper-based books, many people bemoan the coming loss of the book as cultural artifact. In this project we deliberately keep the affordances of paper books while adding electronic augmentation. The Listen Reader combines the look and feel of a real book — a beautiful binding, paper pages and printed images and text — with the rich, evocative quality of a movie soundtrack. The book's multi-layered interactive soundtrack consists of music and sound effects. Electric field sensors located in the book binding sense the proximity of the reader's hands and control audio parameters, while RFID tags embedded in each page allow fast, robust page identification.

Three different Listen Readers were built as part of a six-month museum exhibit, with more than 350,000 visitors. This paper discusses design, implementation, and lessons learned through the iterative design process, observation, and visitor interviews.

Keywords

Audio books, interactive museum exhibits, electronic books, augmented reality, augmented books, multimodal i/o, new genres, sound design, interactive audio, interactive books, page detection, RFID tags, embedded tags, gestural input, smart documents.

PROBLEM STATEMENT: ELECTRONIC BOOKS

Electronic books have been the focus of an enormous amount of research, both commercial and academic. Prior art ranges from Alan Kay's DynaBook to the current crop of commercial CD-ROM publishers and PDA and e-book manufacturers. None, however, have taken the approach of enabling complex electronic interaction through the physical use of a real book. These electronic books are generally presented as an onscreen computer application, as in a CD-ROM, or as a handheld device with a computer screen. The electronic book is almost never conceived as a real book, with real paper pages. Even in the case of "electric paper" the so-called paper is unlikely to offer the look and feel of paper constructed of wood chips, nor will the book much resemble the traditional bound book.

Currently, electronically augmented books tend towards three models: a) button-based interactive children's storybooks, with embedded audio buttons; b) online or CD-ROM books and storybooks; and c) the handheld e-book or PDA-style electronic reader, generally designed for downloadable texts. (As of this writing, a real e-paper book is not available.)

The first instance, where the electronic book is a real book, is children's audio storybooks, which often have short sounds associated with particular images. The sound in these interactive books, in particular those with buttons built into the pages or bindings, is simplistic and often distracts the reader rather than enhancing or illustrating the reading material. However, nearly all of them are designed in a fashion that interrupts the reader's attention to the story by forcing a search for the "hot spots" in the images or text -- the trigger points for the sound, which requires the activation of a button or switch to play and pause. In addition, the sound itself is usually very poor quality, played in short, harsh bursts. The frequency content of such bursts is such that it is difficult for the human perceptual system to successfully put them in the background; this too can disrupt the reader's involvement with the story.

Online or CD-ROM-based "books" lack many of the attributes of books. They retain a book-like depth of content and sometimes a book design interaction metaphor, using pages and chapters, for example; and publishers of CD-ROM books talk about their product as books. But these "books" must be read on a computer, and the distinction

between a CD-ROM or online "book" and other computer applications is unclear. In this case the term "electronic book" is stretched thin, and clearly has more to do with the structure of content than with form or physical affordances.

The e-book or PDA handheld device model of the electronic book has gained a great hold on the imagination of the publishing industry. Though several companies are now marketing these devices, the e-book model of the future of the book has spurred many complaints. [8] In sum: e-books are expensive hardware, they are not sufficiently robust, they have poor resolution or contrast or font formation or there is some other difficulty which makes the screen hard to read for very long. The interface is too present; it requires too much attention; downloading is too difficult and requires extra equipment. Readers find themselves aware of the device, rather than the content; thus immersive reading is difficult if not impossible. Though many of these issues should ease with time and developments in both technology and social expectations around reading, some large questions remain: do people really want to do all their reading on one single device? Is the form and feel of a book a major affordance in focusing the attention of the person reading it?

The Listen Reader draws from research in ubiquitous computing, tangible media, and augmented reality. [3, 11, 12] It pushes back on the concepts of the convergent reading device (as in e-books), of book-as-application (CD-ROMs) or book-as-keyboard (button-based audio storybooks). The Listen Reader is designed to preserve the cherished experience of immersive reading and to preserve the beauty of the paper-based book as token object.

Earlier work: the SIT Book

The Listen Reader is the fruition of several years of research focused on new document genres and the multimodal possibilities of interactive reading, especially. We reported on the SIT (Sound-Image-Text) Book, an early version of this work, in a short paper at CHI'99. [2] With the SIT Book we were primarily interested in the use of audio as an added layer of illustration for a children's book. We also were interested in applying audio as a peripheral layer, added information that was nice to have but that was not necessary to understanding the book. In contrast, the Listen Reader's sounds are an integral part of the interactive reading experience.

Context: museum setting

The Listen Readers are part of a museum exhibition featuring eleven interactive experiences designed to explore the relationship between technology and reading. The six-month exhibit *XFR: Experiments in the Future of Reading* was installed at the Tech Museum of Innovation in San Jose, California. More than 350,000 visitors came through the exhibit between March and September 2000. Because the exhibit took place within the context of a modern technology museum, the XFR exhibits are primarily interactive and hands-on, in keeping with many of the other exhibits.

Fig. 1. The Listen Reader provides a classic immersive reading environment: a comfortable chair, a polished hardwood reading stand, and beautiful paper pages in a soft leather book binder. High-quality embedded audio and responsive proximity ("magic") sensors add to the sense of immersion.

Description

The Listen Reader consists of a large comfortable wingback chair with a small but excellent sound system embedded in it: a small speaker in each wing, not far from a person's ear, and a subwoofer under the seat. A wooden swing arm table sits to the right of the chair; on this table is affixed a book with leather and brass binding and hand-sewn, colorful paper pages. Embedded in the pages and binding of the book – but invisible to the user – are the sensors that enable the electronic augmentation of the book.

The Listen Reader combines an interactive multi-track audio environment with text and images printed on traditional paper pages. Ambient sounds added to each page heighten the drama of the story. As the reader casually fingers different parts of the page, the sound gently adapts to the different moods and scenes. The sound environment is continuous; it stops only when the reader finishes the book, turning the last page.

Here are some examples of typical interactions with a Listen Reader:

A museum visitor -- in this case, a 12-year-old girl -- approaches the three Listen Readers sitting in a calm corner of the museum. Curious at finding comfortable stuffed chairs in this high-tech environment and attracted by the brightly colored books, she sinks into one of the chairs and pulls the book towards her. Suddenly, in response to her hand on the book's cover, she hears a cat's purr. The sounds come from the chair she's sitting in; her head swings around for a moment, trying to locate the source of the sounds, but then she is drawn back to the book. As she opens it to the first page, a rich soundscape fades in: a burbling stream, buzzing cicadas, chirping birds, and an excerpt from

John Coltrane's version of "Stairway to the Stars." At this point the girl realizes she has encountered something new, and she pauses a moment to read the instructional signage mounted on the table next to the book.

"Move your hands all around the book – you don't even have to touch the pages" she is advised, and as she progresses through the book, reading the story on each page, she plays with the sounds by moving her hands, exploring the different atmospheres she can create. As she turns each page, the current soundscape crossfades into the soundscape for the next page; she is never left sitting in silence, not is there an abrupt, harsh transition. Typically, she stays to finish the entire book, and then moves on to read the next Listen Reader.

Another visitor clearly has had much experience with the audio-button storybooks; he tries punching various illustrations to elicit sounds, and is confused when the response is not an instant on/off but rather a volume fade-up or down. Eventually he figures out that he has a different type of control over this sound, and gives up on punching the illustrations. Though he never looks at the instructional signage and does not figure out how to use all of the interaction, he stays to read through the book, content with the shifting soundscapes he gets from simply turning the pages.

A mother sits down in the Listen Reader and piles her three kids around her on the arms of the chair and on her lap. She reads the story aloud to them as they all play with the sounds in the book. Happy to have found a group experience for all of them (and relieved to be off their feet) they read the book all the way through.

MULTIMODAL INTERACTION DESIGN

Listen Reader is an experiment in multimodal reading. In addition to rich graphics and imaginative text, the modality of sound is woven into the storyline. We are interested in sound as well as graphics as part of a multimodal symbol set for reading. Sound as well as graphics can offer strongly affective imagery. Sound also serves to establish a sense of place. Thus, sound can operate as its own additional stylized symbol system, to be read along with the graphics and text. [4]

In the Listen Reader, another point of the audio is to enhance the feeling of communing with a private world, one that gently acknowledges one's own presence. This is similar to the sense of immersion one feels in reading a good book. Combining the two provides an even more immersive experience.

Continuous control versus buttons

An important feature of the Listen Reader is its design for integrating the audio illustration into the story. Interactive children's books often encourage the reader to turn to a new page and immediately punch every image to see whether it makes a noise. We move away from the binary button methodology by giving the reader the soundscape for each page immediately, and then providing the reader with continuous control mechanisms. The task is no longer to find the sounds; it is rather to mix them is a satisfying way, as an accompaniment to the story.

The continuous controllers were mapped to relate the volume of a sound to the proximity of a hand (or other body part). As the reader's hand approaches a certain area of the page, the sound gets louder. It fades when the hand is removed. The continuous controller is tuned to react at the same apparent rate of speed as the hand's motion. The continuous interaction is less likely to induce the seek-and-punch behaviors common in button-based systems. We feel that this allows an important shift in focus for the interactive reader.

SOUND DESIGN AND CONTENT

We chose to begin work with childrens' books, a type of literature where images and text tend to have an almost equal relationship. The three Listen Readers were all children's books from San Francisco's Chronicle Books, for which we designed unique soundtracks. The books were *Hipcat,* by Jonathan London, illustrated by Woodleigh Hubbard; *Armadillo Ray*, by John Beifuss, illustrated by Peggy Turley (we used the Spanish language version, translated by Augustin Antreasayan); and *Frank was a Monster Who Wanted To Dance*, by Keith Graves. Music for *HipCat* was from several John Coltrane albums, used with permission from Atlantic Records; we designed the sound effects.

Each book had a relatively small amount of text per page, and full-page color illustrations ranging from natural settings (deserts, streams, the ocean) to depictions of characters and activities (dancing, eating, driving, reading cat poetry, applauding). The sounds available on each page were related to the images found there. When the page is opened, all the sounds for that page begin to run softly in a continuous loop. There are four sounds associated with each page, one in each quadrant, and each loop is of a different length; thus the effect is that of a changing sonic landscape, not a repeating sound loop.

We chose a running background ambience so that the reader would never fall out of the immersive atmosphere created by the sound. Silence, especially abrupt silence upon a page turn, seems to require the visitor to do something. In the Listen Readers, focus is designed to remain on the story and the multimodal reading experience, rather than on the interaction.

SYSTEM DESIGN

The Listen Reader system combines two technologies: embedded RFID tagging for page identification, and electric field sensing to read proximity data from the hands of the person(s) reading the book. From these two data streams we can establish which sets of sounds are to be played, and what kind of manipulation should be performed on each sound (usually volume or pitch control, or both).

Fig. 2. RFID transponders (tags). Flexible copper printed onto clear plastic, embedded between two printed sheets to create one page thickness for the Listen Reader books. Each transponder has a unique ID. As of September 2000, tags about half this size are available.

Page identification

Systems that gather accurate page-ID data from paper-based books have generally involved either mechanical switch-based systems that are prone to false reads or optical systems that have particular lighting or visual requirements. Though many systems have been attempted, none are in common use.

Camera-based page recognition systems require line-of-sight placement and particular lighting conditions. There are also page-ID systems that require the reader to perform a specific act, such as press a button, run a pen over a barcode, or pass the page through a reader. However these solutions are not transparent; they interfere with a person's natural interaction with the book. For example, Stifelman showed a paper-based electronic notebook that allowed the user to take notes on real paper [9]. Those notes could be cross-indexed electronically because the page-ID was known; but the user had to make sure the edge of the page fit into an optical reader along the side of the notebook. Although it was an elegant solution, it was not completely transparent to the user.

Our method of page-ID uses a new RFID technology (the TIRIS TagIt system from Texas Instruments, with simultaneous ID) to recognize what page a book is open to. [1, 10] A thin flexible transponder tag with a unique ID is embedded in the paper of each page, and a special tag-reader is affixed to the binding of the back of the book.

As the pages turn, the tag-reader notices which tags are within its read range and which have moved out of its range (which is about four inches). The human interacts with the book naturally, and is not required to perform any actions that are not usual in book interaction.

The Listen Reader code (written in Director Lingo, with a specially written serial Xtra) sets the tag-reader so that it continually checks for a list of tags within its reading range.

This list is received by the computer (WinNT 4.0) and updated as frequently as possible (approximately once a second). As a page is turned, the embedded tag leaves the reading range of the tag-reader, and its absence is noticed.

Geometry in tag layout on the pages is of some account. In this system, two tags may overlap by 80% or more, but not completely. This overlap issue is addressable through the development of smaller tags (now available) or multichannel tag readers, or better SID (Simultaneous ID) algorithms.

We use this information to provide appropriate contextual information associated with that page in the book: in this instance, music and sound effects. When a page is turned, the sounds associated with the first page fade down in volume while the sounds associated with the new page fade up in volume.

Other possibilities include spoken word, specific point lighting, ambient room lighting, moving graphics on a computer screen, and scrolling text.

Other tag systems

The Motorola Bistatix tag-based ID system is another example of a thin, flexible tag system that could be used to determine pageturning in realtime. Unlike the Texas Instruments RFID system, which is a passive transponder, the Bistatix is a capacitive field sensing tag, each having its own rewritable ID. This system, though still in development, is perhaps even more useful for our application since it could conceivably be printed directly onto a paper page, rather than having to be embedded between two sheets like the TagIt. Paradiso's swept-frequency tags are another interesting tag technology that could be used in this fashion [5].

Proximity sensing mapped to control parameters

The other technology used in the Listen Readers is electric field sensing. We use it as proximity-sensing technology, embedded in what is essentially a smart binder for the pages of a book. Early versions of this work used an eight-channel sensor called the LazyFish that was developed at the MIT Media Laboratory. [6] The Listen Reader uses a similar commercially available sensor from Quantum Research Labs, the QT110. [7] The sensor measures conductivity within a field. Placing a hand in the sensor's electromagnetic fields will shunt some of the current to ground. The sensor detects the slight current drop, and can track the hand as it passes through the field in all three dimensions. The sensor does not provide a solid x-y-z set of coordinates, as the variables are many (size and shape of person's arm/hand, muscle density, angle of approach, whether they're wearing any metal, etc.). However, it does provide a good real-time set of data that mirrors a user's activity accurately.
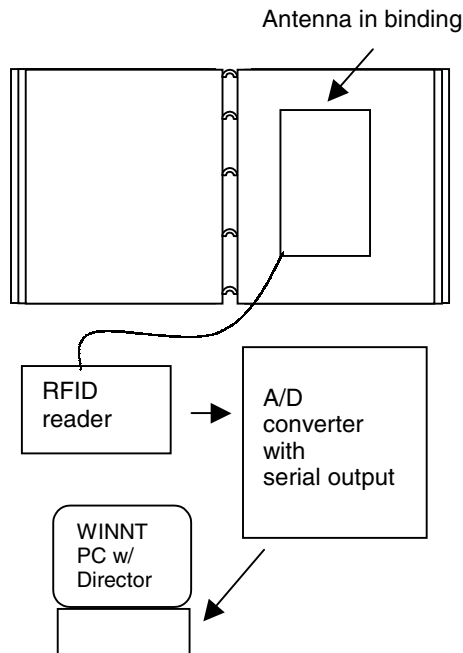
Fig. 3. The RFID antenna is built onto the binding behind right-hand page. As the page is turned, its transponder moves out of the read range of the antenna; the software notices and switches priority to the next page beneath.

In this configuration there are four electrodes in the book binding, one in each of the four corners. This configuration establishes an electromagnetic field corresponding to each quadrant of the book. The sensor is detecting and tracking the reader's hand as it passes through each of the four quadrants.

Physical design

Because of the museum setting, with its huge number of expected visitors (more than 350,000) the Listen Readers were required to be extremely robust, both electronically and in physical design. The pages were a non-tear paper and the book signatures were hand-sewn with a high-strength kite string. The book was attached to a swing arm table, which, along with the chair, was bolted down to a heavy steel plate. All electronics were hidden.

We also wanted to encourage some specific behaviors associated with reading: sitting and spending some time with a book, for example. So, the Listen Readers were designed to provide the look and feel of the comfortable reading corner in one's living room: a cozy chair, good but intimate lighting, beautiful hardwood furniture. The books themselves were hand-bound with brass and leather, and printed on high-quality color printers. The audio was stereo CD-quality, and care was taken to tune each system for premium frequency response (taking into consideration that

the speakers were chair-mounted). By creating a "classy" look and feel, we encouraged visitors to treat the Listen Readers with greater care than if the chairs had been encased in vinyl and the books bolted down with laminated pages.

RESULTS

We were frankly surprised at some of our observations on the Listen Readers, both technical and ideological. Other observations bore out or expanded our original suppositions. The reports below were based on our daily repair and maintenance reports and on observations by our group and museum staff.

Repair, remediation and redesign

Some adjustments were made over the course of the six-month-long exhibition. We developed better systems for replacing the books and their associated files and for rebooting the machines. We fine-tuned the sound content and adjusted the sensitivity of the sensors. As we expected, we had to reset the volume levels to take into account the noisy environment of an active museum. To begin with, we were afraid that the proximity of the three Listen Readers to each other (about fifteen feet apart, set in a triangle) would create cacaphony, with audio from each interfering with the others. However, due to the focus and placement of the speakers within the wings and seat of the armchair, the sound was much louder for the person in the chair than for people standing around watching. In this way the sound added to the personal immersive experience provided by the Listen Reader: a sort of "sound sphere" surrounded people in the chair, an audio atmosphere which rapidly dropped off even two feet away.

We found several different "failure modes" as follows: a) all hardware fine, software frozen; b) software fine, RFID hardware system fine, but the A/D card for the proximity-sensing hardware frozen; or c) all hardware and software apparently fine but the system making no noise. Some occurred once or twice a week, and others occurred only once or twice total (no sound, or RFID frozen). We quickly learned the different types of disk cleanups, reboots and power cycles that should be applied in each case, and designed a script-based daily reboot system that addressed most of these conditions satisfactorily. By midway through the exhibition's run, maintenance on the Listen Readers was routine. In general, they worked; the Listen Readers (like the rest of the XFR exhibit) had less downtime than most of the other museum exhibits.

Robustness

One pleasant surprise was the robust nature of the Listen Readers. We worried that the wooden swing-arm tables might crack under the weight of children leaning on them; that the books would be the target of graffiti artists; that the books were too fragile to withstand daily handling by hundreds of museum goers and that we'd need to replace them every few days. (It turned out be about once every three weeks.) We expected to have to replace or at least

launder the heavy cotton chair covers; but over the course of six months, they withstood the daily onslaught without trauma.

In fact, none of our envisioned disaster scenarios occurred. We believe that this is in part due to the "living room" look and feel of the Listen Readers: beautiful wood, high-quality colorful printed books with real leather binders, soft pooled lighting that highlighted the chair and the person in it. The design of the exhibit itself invited visitors to slow down, sink into the cushions, and enter a calmer state of mind. Our musuem partners also believe that it is in part due to the general character of the visitors to the San Jose Tech Museum; some museums in some areas of the country are notoriously harder on exhibits than others.

Visitor interviews
We observed or interviewed a number of the museum visitors who visited the XFR show. This was done in four ways: informally, through watching and videotaping the visitors' interactions with the Listen Readers; through conversations with the visitors; more formally in an observation report done by museum staff; and a set of ten sit-down interviews with museum visitors who had just finished using the Listen Readers.

Flaunting museum conventions
Received wisdom in the museum world is that visitors do not read. About 30 seconds of time is all that most exhibits can expect to receive, according to the consultants and museum staff we talked with. However, the average time spent at a Listen Reader was much longer, at least several minutes. Frequently visitors would read all the way through one Listen Reader, and then move on to the next. We had hoped that we could encourage people to linger (that's why we built three) but this exceeded our expectations. We believe this is partly because we presented reading as the content of the exhibit, not as signage.

Social reading
One surprise was that the Listen Readers encouraged social reading among the museum visitors. It was at least as common to see two or three or four people (usually children) piled up in one Listen Reader chair as it was to see a single person. Often one person would read the book aloud while the others played with the sounds (this was especially common among family groups, such as a parent with children).

Understanding the interaction
Some visitors were confused by the proximity sensors, expecting to find the kind of on/off functionality of button-based systems. Some such visitors would figure out the difference after reading a few pages, watching other visitors, or reading the instructional signage; others remained frustrated, wanting a clearly causal relationship between the images and the sound, rather than hand motion and sound. We believe that this problem is addressable through a more fine-grained matrix of control (more than

four sensors per page spread). It's also addressable through the complete authoring of books for the Listen Reader system, rather than using existing books to which we add soundtracks.

FUTURE WORK
The *XFR: Experiments in the Future of Reading* museum exhibition, of which the Listen Reader is part, will visit several other museums in the next few years. The lessons learned in each new installation will produce even more robust versions of the Listen Reader's software and hardware. Content may also be redesigned and rewritten; one project is to author a multimodal book for the Listen Reader from start to finish, rather than our current method of adding a soundtrack to existing books. Authoring a Listen Reader from start to finish means we can address the apparent mismatch in causality (what sound is related to which image?) that a few of our visitors mentioned. We can make sure that either the content is authored to each quadrant of the book page, or that the sensors adapt themselves to understand where the important illustrations are upon each page.

Versions of the Listen Reader have applications in several areas other than museum exhibits: for example, language training, musical training, auditory illustration for the blind, and multimodal books for those with ADD (attention deficit disorder). Although the current set of Listen Readers control audio, the same technology could easily drive any kind of dynamic content, such as lights, projections, and so on. For example, a contractor's book of blueprints could control a computer projection or a 3D set of blueprints – by turning to a particular page in a notebook, the contractor could access appropriate and up-to-the-minute drawings, cost estimates, or design samples. We intend to explore some of these applications. We also intend to make a portable version of the Listen Reader, with built-in wireless electronics and a data link to a local area network.

CONCLUSION
Our model of a paper-based electronic book holds promise. People understood its use with far greater alacrity than we had expected, and at least in this application, were delighted to discover such complex functionality in what appeared to be a real book.

The Listen Readers were designed with the idea that form affects meaning, and in fact is inextricable from it. We found that by authoring the form as conscientiously as the content, we were able to achieve some unusual goals: getting people to read deeply in a museum setting, for example, and getting people to read socially, in groups, often aloud to each other. By leveraging people's own knowledge and expectations about the uses of real books, we were able to augment the affordances of the book form with those of an electronic information system.

REFERENCES

1. Back, Maribeth, and J. Cohen. "Page Detection Using Embedded Tags." Proceedings of UIST 2000, ACM Press, 2000.

2. Back, Maribeth, R. Gold and D. Kirsch, "The SIT Book: Audio as Affective Imagery in Interactive Storybooks." Human Factors in Computing Systems; CHI99 Extended Abstracts, 1999.

3. Ishii, H. and Ullmer, B., Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms, in Proceedings of Conference on Human Factors in Computing Systems (CHI '97), (Atlanta, March 1997), ACM Press, pp. 234-241.

4. Mynatt, E., M. Back and R. Want, "Designing Audio Aura." *Proceedings of CHI '98,* ACM Press, 566-573.

5. Paradiso, Joseph and K. Hsiao, "Swept-Frequency, Magnetically-Coupled Resonant Tags for Realtime, Continuous, Multiparameter Control," Human Factors in Computing Systems; CHI99 Extended Abstracts, 1999.

6. Paradiso, J. and N. Gershenfeld, "Musical Applications of Electric Field Sensing." October 1995 *Computer Music Journal*). Also jrs.www.media.mit.edu/people/jrs/lazyfish/

7. Quantum Research Group. White paper on proximity sensors. http://www.qprox.com/whatis.htm

8. Silberman, Steve. "Ex Libris." *Wired,* July 1998.

9. Stifelman, Lisa. "Augmenting Real-World Objects: A Paper-Based Audio Notebook." Proceedings of CHI'96, ACM Press.

10. TIRIS. Tag-it Inlays, www.tiris.com, Product Bulletin, Texas Instruments, 1999.

11. Want, Roy, Ken Fishkin, Beverly Harrison, Anuj Gujar. "Bridging Real and Virtual Worlds with Electronic Tags", ACM SIGCHI. May 1999, Pittsburgh, PA.

12. Weiser, M. "The Computer for the 21st Century." Scientific American, 265(3), 1991, pp. 94-104.