National Library Service for the Blind and Physically Handicapped

The Library of Congress

1291 Taylor Street, NW, Washington, DC 20542; Telephone 202 707-5100; TWX 710-822-1969; FAX 202 707-0712

Technology Assessment and Research Program

A Draft for Discussion

John P. Cookson
Head
Engineering Section

March 27, 1992

Contents

Introduction

The Library of Congress's National Library Service for the Blind and Physically Handicapped (NLS) serves a patron population of about 750,000 with books on special media. Reference (1) describes the genesis, history, and objectives of this service. Reference (2) summarizes the current status of the service, including data on circulation, production, and distribution methodology. The distribution methodology, which is the focus of our technology assessment and research program, and the primary topic of this paper, is further described in reference (3). This description is from a historical perspective and includes current status.

It must be emphasized that service to patrons is the core of the NLS program. All plans for significant changes, which would surely include the introduction of new machines or formats, will, as they have in the past, involve patrons in every step of the process. All innovations will be presented to committees including patrons. Consumer groups will be consulted. All prototype machines will be tested by patrons. NLS will use every means available to assure the products meet the needs of its readers.

For compelling reasons of ready availability, affordability, and compatibility, we are presently committed to cassette tape and flexible-disc distribution media for some years to come (4). These products are affordable, in large part, because they capture the economy of scale inherent in the consumer music market. We recognize, however, there is risk associated with inflexible adherence to any technology driven by this market. The market generally follows the path of improved technology but not necessarily the best technology available.

Other factors such as convenience, price, and glamour appeal are involved in the success of new introductions and the demise of older methods. It is risky business to predict the success or failure of emerging consumer offerings, but it is also risky to be dependent on media and playback machines that may become unaffordable or unavailable because of the underlying technology's fall from popular favor. It is therefore essential that we institute a program to methodically evaluate new technology for possible application to NLS special needs.

This paper presents the elements of such a program, including specific implementation measures. It contains a list of current technologies of interest and describes a conceptual machine that embodies, in a modular fashion, features that might be desirable in a future NLS playback device. Although we place no constraints on what technology may be investigated, our final objective is the specification of new patron access methods and the formulation of strategies to implement them. Reference (5) is our source of basic guidance in visualizing future NLS products that are consistent with our special needs.

The intended audience of this document is program users, all NLS employees, other concerned organizations such as the Research and Development Committee of the National Federation of the Blind (NFB), and agencies in the international community such as the Spanish National Organization for the Blind (ONCE), the Royal National Institute for the Blind (RNIB) in Great Britain, the Canadian National Institute for the Blind (CNIB), and Saudi Arabia's Committee for the Welfare of the Blind. All of these groups have a stake and an interest in the next generation of equipment. More important, everyone has something valuable to contribute. No matter how good new technology may be, it will not succeed unless it first has the universal support of consumers and of the NLS staff. Similarly, the external organizations mentioned above have an advisory, consultative, and perhaps participatory role to play.

The intent of this paper is to suggest a method for adapting technology to NLS needs and for developing a new digital playback machine. This method is structured to involve as many NLS staff as are interested and able to participate. It offers them the opportunity to exercise creativity and inventiveness, with the possibility of being recognized for notable achievements. It distributes a very complex information-evaluation task that might not be as effectively handled by a small cadre. It also recognizes the need for involving external organizations by soliciting their comments and suggestions and by encouraging their participation.

A successful NLS Research and Development (R&D) program will recognize that patron-access strategies are only one component of the very complex product development and delivery system that is embodied in the name NLS.

The primary components of the system are the audio literature on special media and the means to reproduce it--the books, magazines, and machines. The basic product is the audio material, and access strategies include all possible features of a reproducer.

Besides patron access, this system includes mechanisms for managing a myriad of other requirements such as archiving, inventory control, copyright protection, efficient duplication, contract management, and international sharing of material. As it applies to patron access, a successful NLS R&D program will ultimately culminate in the deployment of uncomplicated and reliable digital playback machines.

Books and appropriate playback machines will be the patrons' components of the total NLS system. These will have all the desirable features that are economically and technically feasible. Before that final objective is achieved, however, a successful program will continuously demonstrate to patrons and other interested persons the NLS commitment to quality service through technical excellence. This objective will be achieved by maintaining attractive laboratory facilities with working prototypes available for visitor observation. It will also be achieved by the publication of R&D results in scholarly journals and the presentation of interim findings at appropriate conferences.

A necessary part of the program likewise includes the acquisition and maintenance of local in-depth expertise and hands-on experience with technologies of interest. It also requires the careful formulation of a transition strategy to guarantee uninterrupted quality service during a move from analog to digital technology.

In establishing this formal technology assessment and research program, we recognize the essential place user needs, preferences, and testing play in the ultimate products. We understand the requirement to position consumer assessment of specific technologies at the beginning of the program. Their evaluations will be major driving forces, both initially and when advanced development results in demonstration prototypes. Thus, one of the major components of our program will be the generation of methods to productively involve users throughout the development process. Similarly, methods to involve our international colleagues and other interested groups will be also be pursued.

Program Staffing

The NLS Technology Assessment and Research Program will be coordinated by a program manager who may be any designated NLS employee. Participants will be organized on a technology or research project basis: one principal investigator and one or more coinvestigators for each major technology area or research project. Investigators and coinvestigators may be interested persons from any NLS section or collaborators with external affiliations. The assignment of principal investigator and coinvestigator status will be negotiated among persons expressing an interest in the various technology and research areas.

Participation at any level will be possible only with the concurrence of the proposed participant's supervisor and the program manager. The supervisor must be able and willing to invest the necessary human resources, and the program manager must be confident that the proposed participant's skills match the intended contribution. Participants may have principal investigator or coinvestigator status in more than one technology area or research project.

Funding

The NLS Technology Assessment and Research Program in-house staffing will be funded through the existing NLS payroll. Program hardware and software will be funded primarily by the NLS engineering section budget; however, funding from other section budgets may also occur. Collaborators with affiliations outside of NLS are responsible for providing their own staff, hardware, and software resources. Additional required resources will be requested from NLS management as needed.

The Role of Principal Investigator

Expertise

The first responsibility of a principal investigator (PI) is to acquire in-depth understanding and expertise in the chosen area of interest. In-depth information includes what is available from nontechnical sources such as newspapers and extends to technical journals. For a given technology, such as digital compact cassette (DCC), it includes market research results, the theory of operation, and technical details of implementation. The depth of knowledge must be sufficient to determine what level of resources should be expended in attempting an NLS study and possible adaptation.

Experimentation and Evaluation

The PI briefly formulates, in writing, proposed experiments and evaluation plans. Such plans must include specific milestone dates and cost estimates of required hardware and software acquisitions. Experiments, where feasible, should culminate in a deliverable product. This product may consist of a published report, a demonstration test set suitable for visitor inspection, or both. The PI prepares relevant procurement documents and monitors their status.

For example, a PI in the area of optical media may acquire a compact disk interactive (CDI) unit, demonstrate conventional operation, and then attempt to adapt it in several ways for feasibility exploration. The adaptations might consist of modifying the system software (provided access is granted by the manufacturer) for different audio compression ratios and output rates. An attempt to add features, such as speech synthesis and booksmarks, developed by other PIs, might also be undertaken.

The PI has the option of working alone or may have the assistance of one or more coinvestigators.

Being a PI does not necessarily mean an experiment or project must be undertaken. In some areas, such as in speech-compression assessment, hands-on hardware work is appropriate. In other areas, such as cable TV or direct satellite broadcast, literature research only is currently required. In either case, the essential element is being knowledgeable and conversant in the area of interest.

Reporting

The PI reports orally to the program manager and interested participants once each quarter. If the PI prefers, written reports may be provided in lieu of or in addition to oral reports. Reports summarize relevant events in the commercial market, industry R&D, and NLS investigations.

These reports need not be in a meeting forum. Individual on-site interviews by the program manager may be sufficient. The reports will be recapped in writing, on a biannual basis, beginning in February 1993.

The Role of Coinvestigator

The coinvestigator (CI) assists, advises, and backs up the principal investigator. The CI may negotiate full responsibility for a portion of a technology area or a project and assumes project management responsibility in the absence of the PI.

The CI must have ready access to the corresponding PI's knowledge base and will normally coauthor all reports and publications. The CI relationship makes it possible for persons with a keen interest in a particular field, but no desire to do intricate bench work or other specialized tasks, to manage an area or project. Such a person would assume a PI role and seek a collaborative CI with common interests.

The Role of Outside Collaborators

The interest of collaborators having exterior affiliations will be solicited. These collaborators may come from academia, industry, or other government agencies. We offer collaborators the opportunity to contribute to a socially and technically significant program; we provide an additional application for technology they may develop primarily for another purpose.

For example, certain components of the National Institutes of Health may have an interest in speech recognition for recording radiologist's findings. We may wish to explore speech recognition for possible application to patron interfacing. Another example: the National Science Foundation and its industry partners have an interest in efficient Internet utilization, while we may want to explore the use of Internet for international material interchange.

Collaborators and NLS share findings and results based on research accomplished with their own resources. Public domain computer software will be freely exchanged, assets may be furnished on a loan or transfer basis, and collaborative papers may be published.

Similarly, research results and software developed at NLS will be freely shared with our international colleagues and other interested groups.

The Role of Program Manager

The program manager (PM) is responsible to NLS management for overall administration and coordination of the program. The PM solicits participation, requests resources, conducts quarterly meetings or on-site interviews, prepares written reports and manages external liaison. The PM may also participate as a principal investigator and/or coinvestigator.

The Role of Section Representative

Each NLS section should consider identifying a program representative. This person will monitor the program's progress and develop a section transition plan. The plan will describe, in detail, how the particular section intends to move from current analog technology to future digital technology, without interruption of quality service.

The Conceptual Machine

This section of our Technology Assessment and Research Program Plan contains an idealized description of the next-generation patron-access device. We call this device the Digital Talking Book Machine (DTBM) and explain the reasons for the "digital" descriptor below. The idealization is a conceptual construct that will serve as a point of discussion or reference to help define what technologies may be most attractive, what features may be required, and what features may be optional.

We are not suggesting that this machine can be built at an affordable price with 1992 technology; however, it can serve as a model to aim for and a yardstick to help measure success. We will strive for realistic thinking balanced by guarded optimism. Our guarded optimism will extend to technological progress, funding, copyright regulations, and patron acceptance. The machine description will consist of two parts: general characteristics and specific modules. See Figure 1.

DTBM General Characteristics

Modularity

The DTBM will consist of multiple integrated subsections that are identified on a functional basis. The advantages of modularity include flexibility (the potential for reacting to market forces without full redesign), expandability (the potential for including features as funding and availability permit), and developmental partitioning (the ability to allocate development responsibilities to diverse persons or organizations). There are also reliability and maintainability advantages: single points of failure might be minimized (if a module fails, the entire device need not be inoperative) and a malfunction can be isolated by replacing suspect modules.

Reliability

The DTMB will ideally be available to the patron 100 percent of the time. There are at least two ways to achieve this objective. One way is to ruggedize the device so that severe abuse (drop it down a flight of steps or dump coffee into it) will not cause failure. Another way is to build enough machines so that a patron can have a backup: when one fails, it is removed and the backup is brought into service. In any case, a repair-or-replace maintenance strategy must be adopted. Such a policy decision will require a careful cost analysis that is both technology and market dependent. It may not be possible to achieve 100 percent availability at an affordable price. Nevertheless, availability will remain a high priority concern and will receive continuous attention.

Digital Technology

As the name suggests, the next generation DTBM will be digital in control, data storage, and communications. The most compelling reason for selecting digital technology is that the mass music market appears to be moving in that direction. Digital media afford high-fidelity reproduction without perceptible degradation.

The digital approach also has other advantages such as the ability to manipulate and exchange library material in a way that would otherwise be very difficult, if not impossible. It permits virtually unbreakable copyright protection, precision editing, complex indexing, and remote access. Using read after write, it may also be possible to easily and exhaustively verify the accuracy of each book copy.

Machine resources will be managed by the most capable and affordable microprocessor available. The commitment to digital format will require an investment in building NLS software expertise and development capability. It also has the cost of establishing and supporting a software maintenance infrastructure.

Patron Interfacing

The guiding principle in designing user interfaces is that they must take into account capabilities, limitations, and preferences. One way to implement this concept is to use a hierarchical menu system that allows basic functionality at the highest level and more complex interaction at lower levels, for those who wish to explore. For example, a command to "stop" or "go" would do just that, while internally making a note of the date, time, and start point. Later, the patron could "backup" to the previous start point or some intermediate point by indicating, for example, "backup plus 10 (minutes)." The fundamental concept would be the same as today's linear access to a tape. However, for those so inclined, random access would also be possible.

Specific DTBM Modules

The following is a brief functional description of each DTBM module. Partitioning into submodules and enumeration of salient characteristics is also included where possible. In this context, a module is an aggregate of mechanical and electronic components that act together to provide a specific function; it is not simply a computer program (software) partition but might include software. Software modules will be discussed in a separate section.

Machine Supporting Structure

This module holds the DTBM together mechanically while providing protection from various kinds of damage. It includes temperature control, protection from liquid spills, control of mechanical shock, and control of vibration.

Submodules include:

- Backplane - a known industry standard structure for connecting printed circuit boards such as VME or EISA.
- Enclosure - a rugged case that is distinctive and attractive but discourages pilfering. It supports all of the internal components.

Power Supply

This module provides or conditions electrical energy to operate the DTBM.

Submodules include:

- Battery - when fully charged, capable of operating the DTBM for at least 24 hours; can be fully recharged in one hour; less than one pound in weight; cannot be damaged by overcharge or short circuit; cannot damage the machine by rupture; environmentally benign.
- AC - regulated switching power supply capable of operating the DTBM at specs while recharging the battery; cannot damage the battery; output overvoltage protection; output short circuit protection; input overvoltage protection.

Microprocessor

This module hosts the programs that manage all machine resources. A functional list of applications programs is provided below.

Major features or submodules are:

- Clock frequency of 50 megahertz or faster
- Coprocessor or array processor to support DSP functions such as encryption-/decryption, compression/decompression and data formatting. A neural net processor may also be included to support speech recognition if deemed promising.
- Minimum memory of 32 megabyte ROM and 32 megabyte DRAM (32 megabyte DRAMS are currently in the prototype stage)
- Operating system: real time UNIX - probably Motorola's
- Programming Language: C or C++
- Communications protocols: de facto standards such as TCP/IP and GOSIP
- Data Interfaces: SCSI, RS232, Ethernet, Telephone jack
- Date and time hardware
- Machine status inputs

Microprocessor System and Applications Software

The following is a functional summary of systems and applications software modules that may need to be developed or adapted and integrated. It is assumed that required conventional features found in UNIX are present, such as password protection, remote execute, foreground/background queuing, and multitasking. The presence of a program does not mean that the corresponding feature must be available on the DTBM; software deployment might precede feature implementation or a feature may not be implemented at all.

- System executive (master control)
  Power fail/auto restart
  Menus generator and interpreter
  Scheduler and queue manager

- System peripherals drivers
  Local mass read/write storage
  Audio data compression/decompression
  Optical Character Recognition (OCR)
  Networks drivers
  Diagnostic I/O (keyboard and CRT)

- User interface drivers
  Remote control
  Speech recognition
  Speech synthesis
  Manual controls
  Bookmarks
  Sleeper switch
  Bar code reader (The bar code feature might be useful for tasks such as inventory control, maintenance management, media identification, and media navigation. It is built into Sony's new Lasermax product.)

- Audio output control
  Media recognition and switching (Various media such as DAT, DCC, CDI and various formats such as ASCII, UNIX files, and sampled audio data)
  Decryption
  Reformatting
  Decompression
  Output rate control of audio
  Speech synthesis

- Diagnostics package
  Boot test (memory, CPU, peripherals)
  History of maintenance actions
  Hardware configuration control data
  Software configuration control data

- Usage statistics package
  Date of initialization
  Hours employed for various media
  Hours employed in maintenance mode

Patron Control

This module connects the patron to DTBM resources; it manages user control I/O and evokes responses from other system resources.

Submodules, in order of priority, include:
- Manual controls
- Speech synthesis
- IR Remote control (includes voice input and OCR digitizer)
- Speech recognition
- OCR processor
- Bar code reader
- Sleeper switch

Media Playback Peripherals

- Optical: CD, CDI, CDROM, Mini disk
- Magnetic: DCC, analog Cassette

Local Data Storage

- Magneto Optical disk, 1 Gigabyte
- Mag disk, 1 Gigabyte

NLS audio reproduction equipment has always been limited to "read only." This restriction has been motivated by a variety of factors including a need for economy, simplicity, the safeguard of recorded material, and the discouragement of pilfering. Prototyping the next generation digital machine, however, will require the inclusion of write or record capability. This functionality is needed to conduct experiments in audio-data compression and variable-rate speech. We may also want to experiment with electronic delivery of material. The write capability could be easily and effectively disabled in the final product, at no additional cost, through the use of a software switch. In fact, basic access to the entire machine could be restricted through the use of an identification device such as a password.

Audio Output

- Monaural audio reproduction (D/A converter and amplifier)
- Remotely locatable wireless speaker

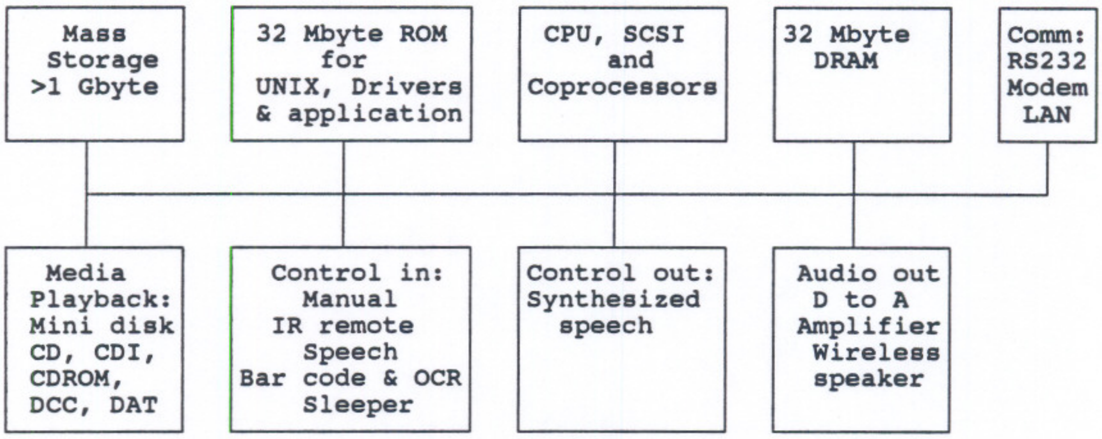

Figure 1. Block Diagram of the Digital Talking Book Machine

Technologies of Interest and Areas of Inquiry

For reasons stated above, in designing the DTBM, our emphasis will be on technology that is likely to appear in the music market. However, we will also pursue other avenues in anticipation of popular adaptation for different mass markets, such as speech recognition for mobile telephones. Access strategies that use electronic distribution rather than circulating media, such as cable TV or Internet file transfer, will likewise be considered. Electronic transfer may ultimately be important in designing international exchange strategies.

In this section we present a prioritized list of technologies that currently appear to have some potential for future NLS use. The list is partitioned into categories corresponding to technology types such as optical or magnetic and it contains preliminary notes on advantages and disadvantages associated with each entry. Positioned within the list are also areas of inquiry that should be researched to enhance program effectiveness. The order in which categories, and items within categories, appear represents an initial assessment of their priority; the most important and promising appear first. The list is dynamic: technologies will be deleted as they fail in the marketplace and will be added as they emerge.

Besides defining the technological topics NLS may wish to explore, the purpose of this list is to stimulate interest and ideas throughout NLS. The program needs participants to become principal investigators and coinvestigators in each of the areas mentioned. Readers may wish to suggest other areas not specifically mentioned where local expertise should be developed. They may also wish to suggest deletion of an item judged to have no possible application to future NLS products.

For several technologies appearing below, adaptation to NLS needs requires access to proprietary design details such as software source code and algorithm parameters. This could present some difficulties. Nondisclosure agreements may be necessary; at worst, access may be firmly denied.

Areas of Inquiry

Productive and efficient involvement of patrons
Transition plans for each section
Conversion of the existing collection to digital format
Conversion of volunteer field production facilities
Field Maintenance of equipment (repair vs replace)
Continuous improvement strategy

Optical Technology

Compact Disk Read Only Memory (CDROM)

+ With compression, perhaps about 28 hours of audio speech capacity
+ Cheap when mass produced, up to 1 gigabyte per disk
+ Popular for software distribution
+ Possible international exchange using UNIX files
+ Caddies keep fingers off and orientation known
+ Rapid random access
+ No loss of fidelity with normal usage
+ Long projected lifetime, perhaps 50 years
- Lack of standards restrict interchangability
- May be difficult to reproduce media rapidly in the field

Compact Disk (CD)

+ Currently the largest selling music medium
+ With compression, perhaps about 19 hours speech capacity (not implemented)
+ Rapid random access
+ No loss of fidelity with normal usage
+ Long projected lifetime, perhaps 50 years
+ Cheap in large quantity; less than $1
+ Home recording possibilities emerging, perhaps at about $5 for 1 copy
+ Built in copy management satisfies music copyright holders
- Standard format allows only 72 minutes of stereo
- There is no way to distinguish between label and information sides without touching
- Fingerprints and other foreign material cause reading problems
- May be difficult to reproduce media rapidly in the field

Compact Disk Interactive (CDI or CDTV)

+ Rapid random access
+ No loss of fidelity with normal usage
+ Long projected lifetime, perhaps 50 years
+ 19 hours of compressed speech implemented
+ Phillips unit at Sears for $799; Similar Commodore CDTV available for $799
+ Sony unit with bar code reader & RS232 also available for $795
+ General purpose potential: 68070 processor, 1 Meg RAM, 512K ROM, UNIX
- Video oriented with unnecessary expensive features
- Same media handling problems as CD
- May be difficult to reproduce media rapidly in the field
- Compression implementation may not be easily alterable

Sony Mini Disk (MD) - Uses optical medium for prerecorded titles and magneto optical (MO) for user recording

+ A protective case is part of the medium
+ Recordability designed in - use magneto optical (MO) medium for recording
+ 74 minutes of compressed stereo music capacity may allow 6 hours of speech
+ Look-ahead memory eliminates tracking errors induced by mechanical shock
+ Built in data compression
+ Rapid random access
+ No loss of fidelity with normal usage
- Standard format allows only 74 minutes of stereo
- Two different media (MO and optical) both unknown lifetimes
- Media lifetimes may differ, say, 10 years for MO and 50 for optical
- Media lifetime unknown but probably at least 10 years
- May be difficult to reproduce media rapidly in the field
- Compression implementation may not be easily alterable
- No firm introduction date; estimated to be during the summer of 1992

<u>Magnetic Technology</u>

Digital Compact Cassette (DCC)

+ Compatible with conventional cassette, similar form factor
+ Minimum effect on patron acceptance
+ Easy to handle, only one way to insert
+ May fit modified mailing container
+ Data compression built in
+ 90 min compressed stereo music capacity may allow 15 hours speech
+ Media cost will probably be less than optical, at least initially
- Introduction scheduled for 3rd 1/4 '92 but may be delayed
- Introduction price high, about $750
- May be difficult to reproduce media rapidly in the field
- Relatively slow serial access
- Media wears with each use but degradation of audio is not gradual

Digital Audio Tape (DAT)

+ Compact, easy to handle
+ Available
- Designed for about 2 hours of uncompressed stereo music
- No provisions for compression
- Expensive
- May be difficult to reproduce rapidly in the field
- Relatively slow serial access

Magneto Optical Disk (MO)

+ Rewritable CD, high capacity (1 gigabyte)
- No widely accepted format standards
- May be difficult to reproduce rapidly in the field
- Lifetime probably less than optical

<u>Communications Technologies</u>

Internet file transfer

+ Funded by Congress via NSF, 1 Gigabit/sec backbone in FY 2000
+ Worldwide government, academic and industry R&D connection
+ Well defined protocols: FTP, UNIX, ISO's OSI and GOSIP
+ Potential international digital exchange standard
- Bandwidth severely limited by local access and loading
- Current bandwidth can't support NLS needs

Telephone connections

+ Most likely channel available to patrons
+ May have potential for interactive use such as catalog browsing
+ A&T encouraging 622 Mbit/s fiber installation - supports book delivery in 5 sec
- Wire is probably too slow to routinely access large data sets

Cable TV

+ 67 percent of US households are wired
+ Has unused bandwidth that is growing with the introduction of video compression
- Few patrons subscribe
- Available bandwidth may be inadequate to be practical

Direct satellite broadcast of TV and other info

+ Antennas and ground stations becoming cheaper
- Few patrons subscribe
- Available bandwidth may be inadequate to be practical

Direct satellite for personal communications

+ May become widely used for personal communications
- Toll grade bandwidth (200 Hz to 3 KHz)
- Available bandwidth may be inadequate to be practical

Digital audio broadcast

+ Motivates intensive development of psychoacoustic compression
- Except for compression, not directly applicable to NLS needs

Speech Processing Technologies

Variable rate audio output

+ User adjustable output rate
+ Digital format allows decimation or replication ann band shifting
- Unproven utility and implementation, needs research

Speech synthesis

+ Good for user interaction, particularly with voice recognition
+ Allows very compact storage (ASCII)
+ May be suitable for informational material, eg. cook book
+ Good quality getting cheaper and more compact (portable translator)
- No artistic value, no inflection
- Good quality is currently expensive (DECtalk)

Limited vocabulary natural speech output

+ Relatively easy to understand, sounds realistic
- Memory intensive

Speech recognition

+ Could permit hands-off remote control
+ Silicon implementations emerging (Odin & Motorola)
+ Intensive development motivated by vehicular telephone application
+ JDR Microdevices (San Jose CA) offers an experimental PC board unit for $130
+ JDR unit could control Sony Lasermax for demo (requires writing demo software)
+ Speaker dependent systems are unlikely to respond to noise or "alien" voices
- Most practical systems are speaker dependent (must be "trained" for each speaker)
- Compute and memory intensive, probably require dedicated processor
- Existing implementations are not practical for a DTBM - expensive, bulky
- Accuracy may not be sufficient

Other Technologies

Advanced Batteries

+ Improved energy density, charge rate lifetime, and environmental acceptability
- Increased cost, restricted availability

Advanced manual controls

+ Capacitive buttons have no moving parts
+ Other non mechanical switches may boost reliability
- Switch identification requires touch which may cause unintentional operation

Remote speaker

+ More flexibility and functionality for the patron
- Increases complexity and unreliability

Remote controls

+ Convenient way to manage playback
+ Might be adapted for verbal command input and OCR input
- Conventional buttons are too small and close together to be useful
- Units are fragile and easily lost

Bar Coding

+ Could help with machine inventory control and maintenance management
+ Could also be used for media identification and media navigation
+ Widely used in libraries
+ In Sony's new Lasermax but standard function is limited to media navigation
- Many incompatible systems in use
- Requires input sensor and software

Full text search

+ May be good for ASCII coded informational material, eg. dictionary
+ Commercial applications emerging, e.g. encyclopedia
- Cannot be conveniently used to search audio sample data
- Compute intensive, may not be practical for very large unindexed data sets

Books on tape

+ Popular products can serve as a model for NLS products
- Their sales value raises concerns about NLS copyright protection

Optical Tape for archiving

+ Vast digital capacity - terabyte (10E12 bytes) per 12 inch reel
+ Rapid access - 60 sec to any file, projected 50 year archive
- Expensive ($250,000 for recorder)

Encryption and Decryption

+ Copyright protection
+ Possible on a per copy basis
- Compute intensive

Optical Character Recognition (OCR)

+ Makes print directly accessible
- Expensive
- Error prone
- No artistic value
- Practical systems currently require manual interaction (page turning)

Sleeper switch with bookmark

+ Marks the book and stops narrating when the reader falls asleep
- Detecting sleep may not be practical or affordable

Proposed NLS Research Projects

In this section we suggest specific research projects that can be immediately undertaken by NLS. The suggestions are presented to stimulate interest and convey a vision of the general content of the overall program. More comprehensive planning for the suggested activities will be the responsibility of the principal investigator. The suggestions are presented in order of priority; what is believed to be most important appears first.

Speech compression. Speech compression is a fundamental issue inherent in the use of digital technology that must be addressed before choices among media can be properly made. We need to determine the relationship between capacity and acceptable quality or between time and required bandwidth.

Projects: Determine the compression algorithm and maximum compression ratio most suitable for NLS purposes. Candidate algorithms might include those employed in DCC and mini-disk. Acquire implementations and modify them to allow variation of compression ratio. Design experiments testing algorithm and compression ratio. Share results by publishing in a scholarly journal. Write functional and implementation specifications.

Variable rate audio output. Variable rate audio is an attractive feature that has been tried before with limited success. The flexibility inherent in the digital domain presents the opportunity to develop it to the point where it is truly practical and useful.

Projects: Acquire a general-purpose digital audio experimentation facility. Test various algorithms to vary the audio speech output rate such as decimation or replication and proportionate increase or decrease of D to A frequency while band shifting as needed. Determine performance parameters and computational requirements. Assess the feasibility of NLS use. Publish results; write functional and implementation specifications.

Internet interchange. Determine the feasibility of utilizing the international Internet infrastructure for the sharing of library materials. Use reference [6] for guidance.

Projects: Find a foreign collaborator with suitable facilities, i.e., access to a UNIX computer on the Internet. Determine the capacity requirements for practical communication of a typical digitized audio book. Obtain, as closely as possible, matching facilities from other Library of Congress programs and assist the collaborator in obtaining the same. Design and execute an experiment to test feasibility. Publish results such as actual data rates as a function of time of day, format conversion delays and conventional (UNIX) data compression ratio achieved. As an alternative, compare results with UNIX files written on CDROM and sent via express mail.

Speech recognition. Research in speech recognition is a high-risk investment, in that the current state of the art doesn't support direct application to NLS needs. However, the potential payoff is significant: possible hands-off control for users. Thus, it deserves some attention, particularly in monitoring the field for significant technological breakthroughs.

Projects: Acquire the most promising recognition system emerging, such as the Motorola implementation of the Odin neural net (designed for hands-off control of vehicular telephones). Use the most advanced form available, such as PC plug in card or breadboard. Design, implement, and demonstrate a voice menu. Determine the limits of performance parameters such as ambient noise, minimum training requirements, and response time. Design and build an interactive demonstration facility; publish results; write functional and implementation specifications.

Other research projects. Many other research-and-technology assessment projects are possible and desirable, such as experiments in speech synthesis, full-text searching, and limited-vocabulary speech output. Others may emerge in the course of development. Such projects will be initiated by the interested PI.

A Technology Assessment Scenario

Rather than conclude this proposal with an abstract summary, the outline of a likely technology-assessment scenario is presented. It is one concrete example of how the program is intended to function.

If I, for example, were to undertake the role of PI for Sony mini-disc, I would first arrange funding to buy several units, recorded material, and blank media.

I would then recruit some coinvestigator volunteers. The team should include at least one person who is blind and no more than one person from the Engineering Section. As a team, we would plan our project, including the purchase of units plus some music and speech selections and blank media. We would also formulate a list of attributes and features to test in a home environment. This list might include items such as ease of use, reliability, portability, and visual and tactile appearance. Coinvestigators would then be asked to take the units home and use them in a way that approximates the use our current cassette machines experience.

All predetermined areas of interest would be carefully observed. Of particular interest would be experiences not anticipated in our planning process. Meanwhile, I would make contact with Sony in an effort to gain access to the machine's internal software. A software development system would be sought so that changes in playback format could be tested. These changes would include band limiting with increased compression, monophonic reproduction and copyright protection coding. If Sony is cooperative, special NLS media could be written and tested.

Ultimately, the experience of coinvestigators would be consolidated and the results of media modification efforts would be described. An interim report would be prepared that summarizes the strengths and weaknesses of this medium as applied to NLS needs. A plan for further investigations or prototyping would also be included. If the technology appeared particularly promising, we would present our ideas to Engineering so that they can write a specification and, in consultation with various section heads, we would propose a transition plan.

Proposed Beginning Schedule

Start-up of the program requires completion of specific steps. A proposed schedule for accomplishing them is presented below. At the beginning of the program, the schedule also represents a prioritization where ranking corresponds to chronological order.

Target Date	Event	Responsible
15 Apr 92	Approve program plan	Chief of MDD
01 May 92	Distribute the program plan to NLS and others	PM
15 Jul 92	Appoint a section rep and transition planner	Section heads
15 Jul 92	Meet with interested persons - clarify & define	PM
01 Aug 92	Recruit a PI for each major area or project	PM
01 Sep 92	Recruit CIs for individual areas or projects	PIs
01 Nov 92	Review literature & summarize status	PIs
01 Dec 92	Identify & price experimental hardware/software	PIs
01 Jan 93	Summarize technology status in writing	PIs
01 Feb 93	Consolidate status into first combined report	PM
01 Feb 93	Propose a one-year consolidated plan & budget	PM
01 Mar 93	Propose a three-year plan with prototypes	PM
01 Mar 93	Review program objectives and status	Chief of MDD

In addition to the above milestones, regular meetings of all participants will be held at least twice a year. At these meetings, the program manager will report on the status of experimental work, PI's projections of significant technological progress, the status of transition plans, and other topics of interest.

References

1. Cylke, F.K., et al, That All May Read, Library Service for Blind and Physically Handicapped People, The Library of Congress, Washington D.C., 1983.

2. "Facts: Books for Blind and Physically Handicapped Individuals", National Library Service for the Blind and Physically Handicapped, The Library of Congress, Washington DC, 1991.

3. Majeska, M.L., Talking Books: Pioneering and Beyond, The Library of Congress, Washington DC, 1988.

4. Questions for the National Library Service for the Blind and Physically Handicapped from FY 1991 Hearing Questions, January 23, 1990.

5. Paris, H., Rasmussen, L., and Dixon, J., "The US Talking Book System: Current and Future Requirements," presented at the International Federation of Library Associations Conference, Moscow, August 1991.

6. Martin, T.J. and Paris, H.B., "Interlending Audio Books for the Blind: Current Access and a Call for a Common Forum," Resource Sharing and Information Networks, Vol. 7, No. 1, 1991.