Using Talking Lights Illumination Based Communication Networks to Enhance Word Comprehension by Subjects who are Deaf or Hard-of-Hearing

Roderick T. Hinman
E. C. Lupton
Talking Lights, LLC
Boston, MA

Steven B. Leeb
Massachusetts Institute of Technology
Cambridge, MA

Al-Thaddeus Avestruz
Talking Lights, LLC

Robert Gilmore
Donald Paul
Nancy Peterson
Boston Guild for the Hard of Hearing
Boston, MA

Abstract

A new method has been developed to transmit auditory and visual information to subjects who are deaf or hard-of-hearing. Ordinary fluorescent lighting is modulated to carry an assistive data signal throughout a room while causing no flicker or other distracting visual problems. In limited trials with subjects who are deaf or hard-of-hearing, this assistive system, combined with commercial voice recognition software, showed statistically significant improvement in sentence recognition, over audio-only or audio-plus-speech reading stimuli.

Introduction

Several different technologies exist for transmitting assistive information to individuals who are deaf (particularly late deafened) or hard-of-hearing, each with its own advantages and disadvantages. FM radio frequency systems use radio waves to transmit audio information. Other technologies commonly used to transmit assistive communications include infrared (IR) transmission and magnetic loop induction. This paper reports on a new means of transmission: modulated room illumination.

Frequency modulated (FM) radio systems broadcast audio from a microphone at radio frequencies near 72 MHz and 216 MHz. Individuals carry a special radio receiver to tune into the broadcast, the output of this receiver can be used with headphones or a small magnetic induction loop worn around the neck to couple to a hearing aid. The system is simple to set up, and is therefore good for a temporary installation. Because radio waves travel through walls, systems in nearby rooms must use different radio channels to avoid interference. Channel allocation can be a problem in facilities with a large number of rooms, although some manufacturers offer units with up to 40 channels.

Infrared assistive systems utilize special IR emitters located throughout a room, typically in a permanent installation. Audio is modulated onto the IR light and received by detector electronics carried by the user. Again, the receiver can drive headphones or induction loops to present the sound to the user. There are no channel allocation problems since walls are opaque to light, and the system works well in a darkened room. However, multiple emitters are needed to cover a large room, and the system requires dedicated equipment.

Magnetic Loop Induction systems couple to hearing aids directly. That is, a large amplifier drives a coil in response to input from its microphone. This coil, typically installed in the floor or ceiling of a room, generates a time-varying magnetic field. This field induces currents in a small coil within hearing aids, the same coil that is used to receive from a neck loop. The main advantage is that users with hearing aids do not have to buy or borrow any additional equipment. The system is costly (thousands of dollars) and the loop must be laid out very carefully to avoid dead zones and interference. It also does not help those without hearing aids.

Talking Lights™ systems are a new method for transmission of auditory and/or visual information. This system makes use of standard bulbs and light fixtures and is particularly well suited for use with fluorescent lights. High efficiency modern fluorescent lights have a flicker rate in the range 40 kHz to 80 kHz, much more rapidly than the human eye can detect. Information can be encoded in the light by modulating the flicker frequency, creating a dual use of this illumination. A special Talking Lights ballast replaces the ordinary ballast needed to operate a fluorescent lamp, as shown in Figure 1. This ballast ignites and keeps the bulb lit, and also modulates the flicker in response to an input source, either analog (e.g. a microphone) or digital (a computer). Figure 2 depicts an example of the light intensity variations; note that there is a non-zero average value and that the frequency is lower in the middle of the plot. An optical receiver with appropriate circuitry is used to detect and demodulate the encoded information (Leeb, Hovorka, Jackson, & Lupton, 2001; Lupton, Leeb, Hovorka, & Jackson, 2000; Hovorka, Leeb, Jackson, & Lupton, 1999). A benefit of the Talking Lights system is its potential low cost and ease of use since it can utilize standard light fixtures and standard wiring currently installed in a facility, and involves no additional energy costs and requires only the installation of light ballasts with communcation capability. Transportable systems using portable lights are also possible.

A prototype analog system has previously been used to transmit and receive audio information (Hinman, et al., 2002; Leeb, et al., 2000). In that configuration, the system is analogous to an FM radio—instead of modulating the frequency of radio waves, the frequency of the variation in light intensity is modulated. Operation is similar to FM, IR, or induction loop assistive systems; the difference is in the transmission means. Like IR transmission, the broadcast is contained within the room. However, the necessity for broad illumination coverage means that the assistive information is available everywhere within that room.

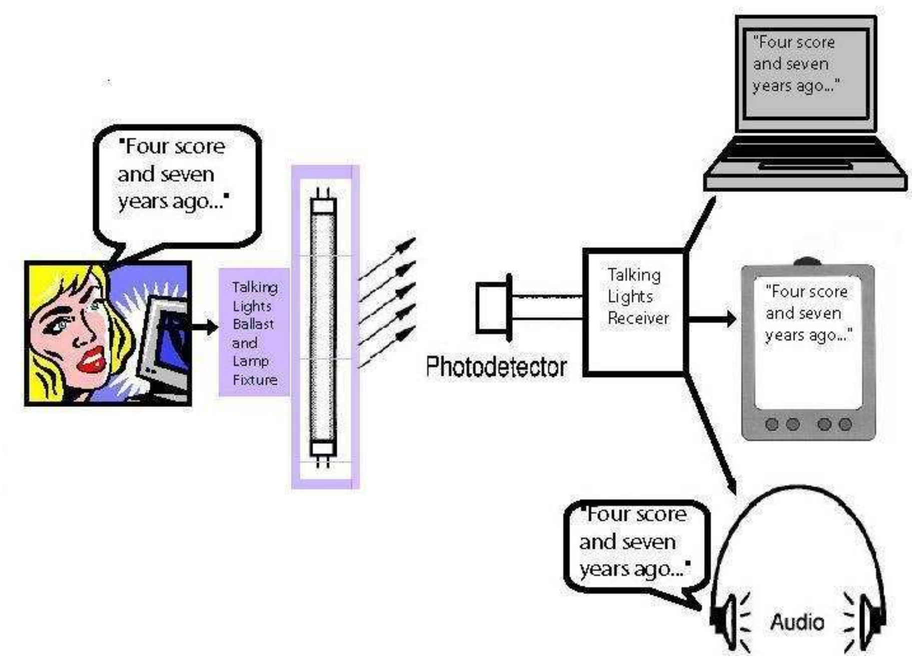

Figure 1: Diagram of Talking Lights concept

Not all users who are deaf or hard-of-hearing are able to benefit from audio information. The ability to transmit text over room illumination is likely to substantially enhance comprehension for these people. The conventional method of captioning is to use a trained court reporter and a communication access realtime translation (CART) captioning machine, to transcribe the spoken words of a speaker in near real-time. These words appear on a large captioning display to assist the deaf and hard-of-hearing with understanding a speech. This system works well, but can be expensive (four to eight thousand dollars of hardware, plus labor costs) and requires the user to sit within reading distance of the display.

An alternative to the large display is multiple individual displays. Every user in the room interested in receiving the text information has a personal data display, which could be a personal digital assistant (PDA, e.g. a “Palm Pilot”-style handheld), a palmtop computer, or a laptop computer. Data from the CART machine would be sent to ballasts with modulation capability. A Talking Lights data receiver anywhere in the room can receive this optical data stream and view the text. The displays could also be easily customizable, allowing text size and other features to be tailored for individual preferences and needs.

A further innovation is to replace the CART system with an automated speech-to-text conversion machine. A number of commercial off-the-shelf (COTS) software packages are available—these run on midrange to high-end personal computers. Conversion accuracy is currently pretty good when the software is trained for the speaker, and acceptable even for a speaker who has not performed the training. Potential advantages of automated speech-to-text include lower costs and greater flexibility in scheduling, as the computer for a lecture hall should always be available. This paper reports on the evaluation of a prototype Talking Lights system to transmit and receive text captioning provided by a live talker and converted by an automated system. The goals were to determine the suitability of the system for the deaf and hard-of-hearing, and measure any improvement in comprehension.

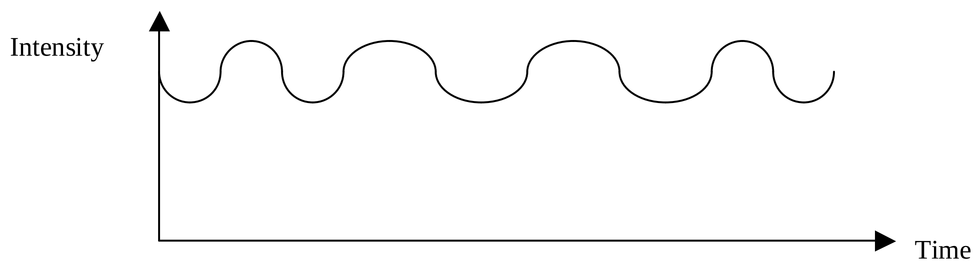

Figure 2: Sketch of light intensity showing frequency modulation of intensity variations

Method

Participants

The participants ranged in age from 24 through 81. Five of the 8 subjects were 69 or older. Five were male and three were female. All participants had hearing loss and supplied audiograms completed during the previous two years. Hearing loss among the subjects ranged from mild to profound. All use hearing aids and assistive listening devices at least part of the time. All subjects could read and speak English although the primary language for two was Portuguese and they considered themselves English language learners (Participants #1 and #2). Characteristics of each person are found in Table 1.

Participant	Age (Yr-Mo)	Gender	Degree Hearing Loss Left	Degree Hearing Loss Right	Speech-reading Skills
1	45-10	F	Profound sensorineural hearingloss. (10/14/99)	Profound sensorineural hearing loss. (10/14/99)	Very good
2	24-10	F	Moderate sloping to profound above 1KHz sensorineural hearing loss. (01/11/99)	Slight sloping to severe sensorineural hearing loss. (01/11/99)	Very good
3	70-11	M	Moderate to severe sensorineural hearing loss. (07/26/95)	Profound sensorineural hearing loss. (07/26/95)	Poor

4	69-8	F	Profound rising to moderately severe from 1KHz to 3KHz then sloping to severe sensorineural hearing loss. (09/05/00)	Profound rising to severe sensorineural hearing loss. (09/05/00).	Below average
5	81-2	M	Sloping mild to profound sensorineural hearing loss. (3/24/99)	Sloping moderate to profound sensorineural hearing loss. 3/24/99	Below average
6	43-1	M	Mild to severe sensorineural hearing loss. (01/12/01)	Moderately-severe to severe sensorineural hearing loss. (01/12/01)	Average
7	80-5	M	Slight sloping to severe sensorineural hearing loss above 2KHz. (08/23/00)	Slight sloping to severe sensorineural hearing loss above 2KHz. (08/23/00)	Average
8	78-2	M	Moderate to severe sensorineural hearing loss sloping to severe above 4KHz. (02/07/01)	Moderately-severe sensorineural hearing loss. (02/07/01)	Poor

Table 1: Participant Data

Stimuli

The eight experimental participants were each given five trials of test sentences. Each trial used one of five variants of assistance. The five variants were:

A	Audio only (no Talking Lights system involved)
A + S	Audio plus Speechreading
A + C	Audio plus Captioning Text
A + S + C	Audio plus Speechreading plus Captioning Text
C	Captioning Text only

Each trial consisted of twelve standard sentences selected from the CID List of Everyday Speech in Davis and Silverman (1970) and read by the talker. The subject was instructed to listen, speechread, or view the captioning as appropriate and repeat the sentence as best they could based upon all available information modes. The talker did not begin the next sentence until the subject had completed their response. The first two sentences were test sentences from the "A" and "B" series for practice. Then the talker would read ten sentences from one complete series, using a randomized series "C" through "H". The response of the subject was monitored and scored by a skilled audiologist. The subject's responses were also audiotape recorded for later review.

Each series consisted of ten sentences that contained fifty "key" words. To score subject performance, the total number of errors was recorded. A "perfect" score would be zero and complete failure to communicate would be 50. The order of the several variants of assistance and the order of the test sentences were randomized for each subject.

Equipment

The complete system is shown schematically in Figure 1. A speaker makes a speech to an audience that includes both people and a personal computer. This transmit side personal computer "listens" to the speech through a microphone, and runs speech recognition software to convert the spoken word into text with high accuracy. The speech recognition software streams this text into a custom piece of software, the Talking Lights voice transcriber, or VSCRIBE. The VSCRIBE software takes the converted text and broadcasts this information as a data stream over the room lights in the test area. Figure 1depicts two kinds of information that could be sent over the lighting network to a variety of devices. In this work, only captioning text was sent to a PC-compatible laptop display device. This laptop runs a second copy of

VSCRIBE, which receives the text stream from the Talking Lights data receiver and displays it in bold text, in real time, as the speech progresses.

Version 4.0 of Dragon System's *Naturally Speaking* software was selected after evaluating several speech recognition engines. *Naturally Speaking* was found to have superb recognition accuracy when trained to work with a particular speaker (typically a 20 minute training period). Interestingly, it was also found to have good accuracy when used with little or no training/customization for a particular speaker. That is, among a group of three male test speakers, the system produced accurate transcriptions (approaching the accuracy of a court reporter) when trained only for one of the speakers. The speech-to-text software interprets the analog voltage waveforms from a speaker's microphone and produces a text output stream corresponding to the spoken word in real time. A Pentium III laptop computer was used as the transmit side computer. The VSCRIBE software running on the transmit side computer intercepts the output stream of the speech recognition engine and sends it out the PC's serial port to the ballast. A copy of the text is also displayed for the speaker in a large text window.

On the receive side, this text window is used to view the converted text transmitted over the optical network. On both transmit and receive sides, a control panel can be used to fully adjust the display characteristics of the text. Font size (between 7 and 48 point), line spacing, and window size can all be selected. Before beginning the trials, each subject was given the opportunity to modify the display parameters. All felt it was readable at the default 24 point size.

The digital data has been encoded in the light in such a way as to eliminate the possibility of visible flicker due to the modulation. A simple data-encoding scheme would use one frequency to represent a "one," and another to represent a "zero." Even though the lamp drive frequency is much greater than the eye can see, it is possible for a long string of digital "ones" followed by a long string of digital "zeroes" to produce flicker in the perceptible range (Buffaloe, Jackson, Leeb, Schlecht, & Leeb, 1997). The internal design of most ballasts makes the lamp's average intensity mildly dependent upon drive frequency. A slightly more complex coding scheme breaks each bit into equal periods of both frequencies, so that any intensity variations occur at twice the bit rate (Leeb, et al., 2001). The bit is encoded by the *order* of the two frequencies within the specified period. This so-called Manchester coding is used in magnetic recording, Ethernet, and other situations where low frequency variations cannot be tolerated.

Procedure

Comprehension experiments were performed in a medium sized conference room. Subjects were seated two meters from a live talker, who read sentences to them and into the microphone of the speech-to-text computer. The receive side computer was placed in front of the participant; during trials without captioning, its receiver electronics were disabled. A recording of the party noise made by about 400 people in Copenhagen, Denmark, was played during all trials via loudspeakers at about 60-65 dBHL (Widex).

The subjects were allowed to use their personal eyeglasses or contact lenses in order to see the talker and the captioning device, although no assessment of visual acuity was made. For the trials A, A + C and C (those without speechreading), the subject's view of the talker's face was blocked. The subjects were also allowed to use their personal hearing aids for all trials, except C - Captioning Text only. In that instance, hearing aids (if any) were removed and sound blocking hearing protection was used to limit reception to visual only.

Results

The eight experimental subjects showed extensive variability in their capability to understand and repeat back the test sentences. Table 2 shows the accuracy of repeating back of the test sentences for each subject and trial, with 100% representing perfect accuracy and 0% representing complete inability to repeat back any key words. In Table 3, the number of errors per trial for each subject are listed. The number of errors per trial ranged from a high of 43/50 to a low of 0/50. Mean accuracy of sentence recognition for the five information conditions are also presented at the bottom of each column in the tables.

Participant	Auditory Alone	Auditory & Speechreading	Auditory & Captioning	Auditory, Speechreading & Captioning	Captioning Alone
1	14%	72%	82%	94%	96%
2	42%	96%	94%	98%	98%
3	44%	84%	100%	98%	100%
4	44%	92%	98%	96%	94%
5	82%	92%	98%	100%	94%
6	96%	100%	100%	98%	100%
7	96%	92%	100%	95%	98%
8	98%	100%	100%	100%	100%
Average over participants	64.4%	91.2%	96.8%	97.4%	97.2%

Table 2: Accuracy of Repeating Back Key words by Subjects for Each Assistive Mode

Participant	Auditory Alone	Auditory & Speechreading	Auditory & Captioning	Auditory, Speechreading & Captioning	Captioning Alone
1	43	14	8	3	3
2	29	2	3	1	1
3	28	7	0	1	0
4	28	4	1	2	3
5	9	4	1	0	3
6	2	0	0	1	0
7	2	4	0	2	1
8	1	0	0	0	0
Average over participants	17.8	4.4	1.6	1.3	1.4

Table 3: Errors Made by Subjects for Each Assistive Mode (50 possible per trial)

The data were analyzed with a one-way within subjects ANOVA (five levels: Auditory only, Auditory with Speechreading, Auditory with Talking Lights Captioning Text, Auditory with Speechreading and Talking Lights Captioning Text, and Talking Lights Captioning Text only). The effect of the different information conditions was significant, $F(4,28)=8.9$, $p<.001$, $MSe=181.2$. As can be seen in Table 3, the effect is largely attributable to poor sentence recognition for the auditory only condition.

However, people with impaired hearing would customarily be able to speechread as well as listen. Comparing that condition (Auditory with Speechreading) to the three conditions containing captioning text shows that having text information available resulted in significantly greater word comprehension. The specific significances are:

a) Auditory with Captioning Text ($F(1,7)=7.13$, $p<.02$);

b) Auditory with Speechreading and Captioning Text ($F(1,7)=5.2$, $p<.03$, and;

c) Captioning Text only ($F(1,7)=4.6$, $p<.04$.

In the analysis, separate within subject error terms were used for each comparison based just on the conditions involved.

The participants in the trial were enthusiastic about the new method and found it a valuable method for them to increase their comprehension. They were particularly encouraged by the projected low cost of the system. One participant commented that the system was a "wonderful way to fill a room with information."

Discussion

Not surprisingly, the results suggest that individuals with moderately severe or greater hearing loss tend to most benefit from the availability of visual information that increases the redundancy of the message and thereby increases the listener's confidence in what they thought they heard. These individuals consistently demonstrated the greatest improvement in sentence scores when comparing auditory only (A) to any of the other modes, those that include visual with or without auditory information (A+S, A+C, A+S+C, and/or C).

Of the two visual modes, captioning text produces higher accuracy than speechreading, even with imperfect speech-to-text translation. Typical errors for the voice recognition engine are homonyms or similar sounding words. Most people were able to filter out any translation errors to arrive at correct responses. However, participant #5 found the captioning only (C) trial difficult, because he had to think about words on the screen and what they might actually be.

Several participants noted that the automatic speech-to-text conversion process did not supply any punctuation, increasing the difficulty of parsing the sentences. The current state of the art in voice recognition is for someone dictating to a computer to speak the punctuation mark (e.g., say the word "comma") at the appropriate place in a document. Also, most recognition engines work best when pretuned for the speaker. These issues complicate real-time translation of a speech given to a live audience, especially situations like a town meeting where there could be many speakers. However, voice recognition technology is rapidly progressing, and both of these issues may be addressed in time. Neither is a limitation if a human CART reporter generates the text to be broadcast over the room illumination.

The uniqueness of providing captioning via illumination is in the transmission and display method, not in the source of the text itself. There will be many situations where a CART reporter is the preferred method of speech-to-text translation. Light-based transmission allows the user to have an individualized display device with features not available with the typical projection system used with CART. Display characteristics like font size and scroll speed can be modified to individual preferences, and the person may sit wherever he or she wishes. One participant liked the system because he could see what had been said even if he completely missed the auditory information. Conversation history is also available on a CART projection system to the extent of the size of the screen. Individual Talking Lights receivers store the text data and so enable the user to review the entire conversation or speech, if they so desire.

The hardware used for this evaluation was a prototype; the custom electronics (ballasts and receivers/decoders) were large and not tested for regulatory and safety compliance. The software ran on laptop computers for ease of development. A smaller, more convenient receive side platform would be greatly desired. Ideally, it would be a PDA with a receiver built in or inserted into an expansion slot.

When these components are developed, they will need to be tested more rigorously than in this preliminary study. Future evaluation should use more challenging test procedures and stimuli to more closely approximate real life listening situations and delineate more clearly any benefits for individuals with mild to moderate hearing loss. Some examples are word lists versus sentence lists so that contextual cues will not play a role in individual performance, and high frequency word lists to increase the challenge for individuals with high-frequency hearing loss.

Conclusion

Illumination can be used as a carrier of assistive information for deaf and hard of hearing individuals. A system which uses ordinary fluorescent lights to transmit this information has been demonstrated. This system provides illumination with no flicker or other visually distracting problems while also transmitting the assistive information. This system was used to transmit text generated by speech-to-text computer software to users. In trials of this system with a limited number of deaf and hard of hearing users, the text assistance resulted in statistically significant improvement in comprehension over auditory only or auditory plus speechreading sstimuli.

Acknowledgment

The authors would like to thank Prof. Randolph Easton of the Psychology Department of Boston College for valuable discussions concerning the statistics of this work. Support for this work was provided by the National Institute on Deafness and Other Communications Disorders under SBIR Grant 1R43DC04015-01.

References

Buffaloe, T., Jackson, D., Leeb, S., Schlecht, M., & Leeb, R. (1997). Fiat Lux: A Fluorescent Lamp Transceiver. *Proceedings of IEEE Applied Power Electronics Conference, Atlanta, GA, 1997.* New York: IEEE Press.

Davis, H., & Silverman, S. (1970). *Hearing & Deafness.* 3rd ed. New York: Holt, Reinart & Winston. CID List of Everyday Speech, pp. 492-495.

Hinman, R., Leeb, S., Avestruz, A., Lupton, E., Bentsen, B., & Easton, R. (2002). *Dual Use Lighting for Assistive Communications*. Paper presented at NSF Design, Service and Manufacturing Grantees and Research Conference, San Juan, PR, 2002 (on CD, no paper version). Ames: Iowa State University Engineering Communications and Marketing Department.

Hovorka, G., Leeb, S., Jackson, D., & Lupton, E. (1999). *Analog and Digital Electronic Transceivers for Dual-Use Wireless Data Networks. PCT Application WO99/53633, October 21, 1999.* Geneva, Switzerland: World Intellectual Property Organization.

Leeb, S., Hovorka, G., Jackson, D., & Lupton, E. (2001). *Dual Use Electronic Transceivers for Wireless Data Networks. US Patent 6,198,230.* Washington, DC: U.S. Patent and Trademark Office.

Leeb, S., Hovorka, G., Lupton, E., Hinman, R., Bentzen, B., Easton, R., & Lashell, L. (2000). *Assistive Communication Systems for Disabled Individuals using Visible Lighting.* Paper presented at the 15th International Conference on Technology and the Disabled, Northridge, CA, 2000.

Lupton, E., Leeb, S., Hovorka, G., & Jackson, D. (2000). *Communications Systems. PCT Application WO00/30415, May 25, 2000.* Geneva, Switzerland: World Intellectual Property Organization.

Widex Hearing Aid Company. *Party Noise*. Long Island City, NY. A recording of the party noise made by about 400 people in Copenhagen Denmark, played at about 60-65 dBHL.