Running head: EFFECTS OF PRINTED VERSION OF THE TEXT ON LEARNING

Effects of Adding a Video Image, Printed Text of a

Lecture, or Both, on Learning of an Auditory Lesson

Alexander Fraser

Algoma University College

March 23, 2004

Abstract

Extensive literature on technology for teaching, which shows inconclusive and sometimes contradictory information about the use of printed text and a visual image of a speaker's face, has been produced. Further research is needed to clarify the role of both subtitled text and a visual image of a speaker has on learning.

Effects of Adding a Video Image, Printed Text of a

Lecture, or Both, on Learning of an Auditory Lesson

Traditionally, teachers using the lecture format address their students face to face. Students can process the information by both hearing the speaker's words and seeing the lecturer speak. Written text, with or without a visual image of the lecturer speaking is an additional way of processing information. There is a body of research (Garza, 1991; McGurk, 1976) studying how each channel of information contributes to learning and retention. Much of this involved the addition of a text stream to an audio or audiovisual presentation.

The practice of combining printed text with audiovisual presentations is not a new one. For centuries, teachers have been writing important points on chalkboards while they say them aloud. Some researchers (Bean & Wilson, 1989; Garza, 1991; Goldman & Goldman, 1988) predict that text subtitling will have a positive influence on learning but others (Spanos & Smith, 1990) suggest that subtitled text may serve as a distraction or block to information processing for some learners. Most research on subtitling, including closed-captioned television, has been aimed at increasing entertainment value, or else has studied populations that somehow have a block or impairment of information processing, such as learners of a second language or hearing- impaired learners.

Same Language Subtitling for Second Language Acquisition and Other Special Cases

Vanderplank (1988) studied the value of text subtitles in language learning by learners of English as a second language. His study group consisted of 15 European and 8 Arabic exchange students studying English language, translation, and interpreting. They varied in English language level from low intermediate to post-proficiency. They viewed

recordings of popular British and American television programs, with subtitles added, for one hour each week for nine weeks. All participants reported finding the subtitles beneficial but two participants thought they would have listened harder, and improved their listening abilities, without text.

Garza (1991) used verbatim captioning with 110 participants including 70 adult learners of English as a second language and 40 adult learners of Russian to explore the benefits of combining spoken words, video imagery, and printed text simultaneously. He presented five short verbatim captioned segments from Russian and English TV programs over a one-month period, and tested students' ability to use specific words from the programs when they recalled the content of the programs. His 2 X 2 between subjects design had independent variables of language (English and Russian) and captions (with captions and without captions). Garza reported that there was significant improvement in global comprehension for those who received captioned video presentations over those who did not, but he did not perform significance tests. Fortunately, he included descriptive statistics, which made it possible to perform significance tests showing that there was a significant main effect of language and of captions with no significant interaction, suggesting the use of captions enhances learners' competence in reading and listening. Garza used participants that already possessed a high level of competence in the subject areas (either English or Russian).

Borras and Lafayette (1994) investigated the effects of subtitling on oral communicative performance. Their participants were 44 university undergraduate studying French. In a 2 X 2 factorial design, the independent variables of subtitling and task level of an oral procedure were manipulated to produce the four treatment

conditions, (n = 11), which were: subtitled video with a lower level task, unsubtitled

video with a lower level task, subtitled video with a higher level task, unsubtitled video

with a higher level task. The dependent variable, oral communicative performance, was

defined as success at conveying information orally as measured by a test of speaking

ability. Compared to the unsubtitled condition, the participants in the subtitled condition

had significantly higher overall oral performance scores. Neither the effect of task level

nor the interaction of task level with the subtitling variable was significant. The authors

concluded that subtitles have potential value in helping learners to better comprehend

linguistic input and to produce comprehensible speech.

Some research was conducted to support the idea that same language subtitling

improved memory for verbal material. Kothari, Takeda, Joshi and Pandey (2002) studied

the impact of same language subtitles on the reading skills of disadvantaged school

children. Three groups of Indian children, (n = 46), were randomly selected from grades

IV and V. The participants spoke various dialects but were all presented with Hindi films

over a three month period. There were three conditions: songs, songs and subtitles,

neither song nor subtitles. All the participants were given a test to measure the

recognition and reading of syllables. There were no significant differences found between

any of the groups. Although there were no definite conclusions the authors felt that

reading skills could improve simply from exposure to the language, even through songs

without subtitles.

Research generally supports the hypothesis that the addition of subtitles

enhances learning of a second language. In addition, there has also been research using

subtitles or closed-captioned television programming to enhance learning for populations

other than second language learners, including those who are hearing impaired, those with low literacy, and those who have other learning disabilities.

Closed-captioned television was first developed to aid the deaf but there has been recent interest in its use by hearing audiences with low literacy levels. Goldman and Goldman (1988) studied the use of closed-captioned television to help remedial readers at the high school level. They taped popular television programs and presented 25-minute programs twice a week. The programs were presented without audio input, forcing the participants to read the text dialogue. This looked at the use of subtitles without the benefit of an audio component. There was no control group or significance test of significance performed. The readers were instructed in various reading skills before, during, and after watching a program. All students scored 70 percent or higher on 20-question comprehension and vocabulary tests. The authors concluded that the closed captioned program was effective in increasing motivation. Students missed fewer captioned classes and paid more attention to the class material than they did in normal class sessions, perhaps because they enjoyed the captioned classes more than the regular classes. The authors stated that students displayed more motivation and interest in learning in general during the closed-captioned television classes than in their other classes.

Bean and Wilson (1989) examined the effects of closed captioned television as a medium for sight vocabulary development. The participants were 24 students in an adult literacy program. All participants scored below 7th grade equivalency on a standardized reading comprehension test. The research used a 2 X 3 between-subjects design. The first independent variable, reading level, had levels of easy and hard. The second independent

variable, learning approach, had experimental conditions of closed captioned television with instruction, closed captioned television without instruction, and script with instruction. All participants were given a word recognition pretest and posttest. The effect of the various approaches on student performance was also rated by a fluency measure after each lesson. Results of the pretest and posttest indicated an overall significant recognition performance improvement. However, student performance did not differ across treatments and all treatments were equally effective. The lack of differences might be due to a ceiling effect. Simple exposure to the test material might explain the increase in performance improvement.

Goldman and Goldman (1988) studied the use of closed-captioned television to help remedial readers at the high school level. They taped popular television programs and presented 25-minute programs twice a week. The programs were presented without audio, forcing the participants to read the text dialogue. The readers were instructed in various reading skills before, during, and after watching a program. All students scored 70 percent or higher on 20- question comprehension and vocabulary tests. The authors concluded that the closed captioned program was most effective in increasing motivation. Students missed fewer captioned classes and paid more attention to the class material than they did in normal class sessions perhaps because they enjoyed the caption classes more than the regular classes. Students displayed more motivation and interest in learning in general during the closed-captioned television classes than in their other classes.

The Effect of Image of a Speaker on Learning

Hearing impaired people often read the lips of speakers to understand speech. For normal hearers the auditory modality is most important for speech perception, but the

visual modality or lip reading allows all listeners to better understand speech (Skhiri,

2001). Lip reading is something people do unconsciously and effortlessly. Listeners who

watch a video wherein the speaker's audio output and visual lip movements match can

accurately repeat what was said but 92 percent of listeners who receive conflicting audio

output and visual lip movements repeat back incorrectly (McGurk & MacDonald, 1976).

This phenomenon is called the McGurk effect. The addition of an image of a lecturer's

face and lip movement to an auditory lesson might allow the listener to receive the

message more accurately and to enhance learning of the lecture material.

Analysis: Why and When Will Additional Information Help

It is not clear why text enhances learning. Several explanations, such as the

entertainment factor associated with watching videos, dual-channel coding, and

individual learning style of the learner have been suggested. These can be analyzed to

explain the effects of adding subtitled text to an auditory lesson.

Perhaps closed captioning makes audio presentations more entertaining because

most learners are familiar with television and associate video presentations with

entertainment. If the material presented is intrinsically entertaining, then any

manipulation that makes the material easier to follow will enhance memory for the

material. In Goldman and Goldman's study (1988) described above, enthusiasm for

watching closed captioned television was so high that a closed-captioned television club,

where students watched a captioned popular television program every day at lunchtime,

was founded. So many students from outside the reading program were attracted to the

club presentations that people had to be turned away unless they had a reservation for a

showing. The allure of watching television, even programming with subtitles, might have

motivated the students to practice reading.

The entertainment factor is not the only thing that might explain the positive effects of adding subtitles to lessons. Another possible explanation centers on how the input material is coded or processed. Paivio's (1990) dual coding theory suggests that verbal and non-verbal processing are two distinctly specialized cognitive subsystems used in processing information. One subsystem processes images that Paivio calls "imagens", while the second subsystem processes what he calls "imogens" or verbal utterances. The underlying assumption of his dual coding theory is that all learners can process information through the two subsystems simultaneously without them interfering with or blocking each other. Perhaps auditory (spoken) and visual (text) presentations of the same message could complement one another in this way.

Other research supports the view that dual coding of audiovisual and subtitled text works especially well for people who have a problem understanding audio input. Spanos and Smith (1990) reported that closed-captioned videos allowed less proficient students to understand individual words from either an audio or video track. Whereas more proficient students were able to process language from both tracks, the less proficient students appeared to use only one, either the audio or video track. Attempting to attend to both tracks might burden cognitive resources of the less proficient students to a point of being a distraction to the processing of the information. This supports the hypothesis that hearing impaired learners, or learners with listening problems, might benefit from having information processed through a visual medium. People who have no trouble understanding auditory information might unconsciously use the medium that supports their individual learning style, but people who do have trouble processing auditory

information would have the opportunity to process the information through an alternative medium.

Learning styles (Learning Styles and Approaches, 2002, para.1 and 2) are commonly divided into three main categories: visual (people who learn best by reading or observing), auditory (people who learn best by hearing or listening), and kinesthetic (people who learn best by doing). People tend to learn predominantly through one of their senses, i.e. auditory, visual or tactile. Auditory learners tend to retain what they hear, visual learners tend to retain more of what they see and read, and tactile learners tend to retain more when they are using their sense of touch. The addition of printed text to an auditory lesson offers learners the opportunity to process the information either auditorially or visually.

Proponents of the use of subtitled videos argue that subtitles may help learners develop language proficiency by enabling them to be conscious of language they might not otherwise understand but opponents of the use of subtitled videos argue that subtitles are distracting and that they slow down learners' listening skill development. Borras and Lafayette (1994) suggest that arguments by both sides are mainly intuitive and lack empirical evidence. Perhaps multiple coding of information provides learners with the option to decode the content of the material with a decoding strategy that best matches their natural learning style, offering more ways for learners to remember the lesson content.

Summary

The addition of subtitles and/or a visual image of a speaker's face to a taped lecture enhance learning. Research generally supports the hypothesis that the addition of

subtitles enhances learning of a second language and that hearing impaired individuals benefit by seeing the speaker's face as he or she talks. It is not clear why text or visual image enhances learning but dual-channel coding, leaning style, and the entertainment factor associated with watching videos are possible explanations.

The advent of distance communication technology raises questions of the costs versus benefits of transmitting various kinds of information. There is considerable time, effort, and expense involved in producing subtitled audiovisual lessons but with technological advances, subtitling is a more accessible option, but at a cost of increased bandwidth. This increase is justified by improved learning.

References

Bean, R.M., & Wilson, R.M. (1989). Using closed-captioned television to teach reading

to adults. *Reading Research Instruction, 28(4)*, 27-37.

Borras, I., & Lafayette, R. C. (1994). Effects of multimedia courseware subtitling on the

speaking performance of college students of French. *The Modern Language*

Journal, 78(1), 61-75.

Garza, T. J. (1991). Evaluating the use of captioned video materials in advanced foreign

language learning. *Foreign Language Annals 24*(3), 239-258.

Goldman, M. & Goldman, S. (1988). Reading with closed captioned TV. *Journal of*

Reading, 3 31(5), 458.

Kothari, B., Takeda, J., Joshi, A., & Pandey, A. (2001). Same language subtitling: A

butterfly for literacy? *International Journal of Lifelong Education, 2*(1), 55-67.

Learning styles and approaches, (2002). Retrieved January 30, 2002, from

http://academic.son.wisc.edu/cnp_orient/OnlineLearning/Learningstyles.htm

McGurk, H. and MacDonald, J. (1976). *Hearing lips and seeing voices*. Nature, 264,

746-748.

Paivio, A. (1990). Mental Representations. New York: Oxford University Press.

Skhiri, M. (2001). Visual cues in speech perception. Retrieved November 19,

2003, from Linkoping University, Swedish National Graduate School of

Language Technology Web site:

http:www.speech.kth.se/~rolf/gslt_papers/MustaphaSkiri.pdf

Spanos, G., & Smith, J. (1990). Closed captioned television for adult LEP literacy

learners. *Electronic Resources Information Center Digest*. Washington, DC:

National Center for ESL Literacy Education. (EDRS No. ED 321 623).

Vanderplank, R. (1998). The value of teletext sub-titles in language learning. *ELT Journal, 42*(4), 272-281.

Running head: EFFECTS OF PRINTED VERSION OF THE TEXT ON

LEARNING

Effects of Adding a Video Image, Printed

Text, or Both, on Learning of an Auditory Lesson

Alexander Fraser

Algoma University College

February 29, 2004

Abstract

The present research measured the effect that printed text, or visual image, or both together had on learning information presented auditorally. The participants were 64 Algoma University College first-year students who were randomly assigned to four groups corresponding to the four conditions. Each group was given a six-minute videotaped lecture and tested immediately. It was predicted that learners who were presented with an auditory lecture that had additions of text, or visual image, or both together would score higher on a learning posttest than those that were presented with the auditory lecture alone. There were no significant differences found between the groups.

In its most basic form, human communication involves speaking and listening. One can communicate or teach using only an auditory channel, and distance education in the past was often carried out with audiotapes of spoken lectures. As technology has advanced, it becomes increasingly practical to add additional information to the auditory stream, though such additions necessarily come at a certain cost. Therefore it becomes useful to investigate the benefits resulting from such additional information, and to see how different additions may interact with one another.

The obvious possible additions to the auditory channel both involve vision: one possibility is to add the visual image of the speaker's face. This enables the learner to benefit from articulatory cues to help understand what is being said. These cues are normally present in face-to-face conversation, and there is clear evidence that they are employed automatically and unconsciously when we hear and watch a speaker (McGurk H. & MacDonald J., 1976). We can understand a speaker better when we see his or her face speaking the words.

The other potential addition is verbatim subtitling: the word-for-word printing of the spoken message. The viewer can then read the message while listening to it. Subtitling has been widely used for foreign-language films, for second language learning and for special populations (e.g. with auditory handicaps); it has not, however, been widely employed in distance education.

It would be of value to quantitatively assess how much (if at all) each of these additions to the auditory message aids learning and retention. It is also of interest to see whether they interact: one could argue that, since both involve visual processing,

presentation of the facial image and the sub-text together would lead to competition; if this occurs, addition of both would offer no advantage over presenting either one alone. On the other hand, given that use of facial cues is automatic and unconscious, one might be able to utilize those cues at the same time one is reading the words of the subtext. This would lead to a processing advantage rather like that described in Paivio's "dual coding" theory, which says that auditory and visual channels support each other rather than competing. The result would be that the combination of visual image and subtitling would be of more benefit than either one alone.

The addition of subtitles to audiovisual presentations to enhance learning has been used predominantly for, or else by, populations who are hearing impaired, learning disabled, or learners of a second language. The present research is designed to explore the impact of subtitles added to an auditory lecture in a more general population.

Garza (1991) used verbatim captioning with 110 participants including 70 adult learners of English as a second language and 40 adult learners of Russian to explore the benefits of combining spoken words, video imagery, and printed text simultaneously. He presented five short verbatim captioned segments from actual Russian and English TV programs, over a one-month period, and tested students' ability to use specific vocabulary from the programs when they recalled the content of the programs. His 2 X 2 between subjects design had independent variables of language (English and Russian) and captions (with captions and without captions). Groups were randomly divided so that half would view with captions (test) segments and half would view the without captions (control) segments. Garza felt there was

significant improvement in global comprehension for those that received captioned video presentations over those that did not but he did not perform significance tests. Fortunately he included descriptive statistics, which made it possible to conclude there was a significant main effect of language and of captions with no significant interaction, suggesting the use of captions enhances learners' competence in reading and listening. It may be important to note that Garza used participants who already processed a high level of competence in the subject areas (either English or Russian). The present research also explores the effects of subtitles on learning but differs from Garza's research in a few ways. The present research uses mainstream learners who are presented with lecture material on a topic with which they may, or may not, be familiar. Garza's research had independent variables of language and subtitled text whereas the present research has independent variables of subtitled text and image.

Hearing impaired people often read the lips of speakers to understand speech. For normal hearers the auditory modality is most important for speech perception, but the visual modality or lip reading allows all listeners to better understand speech (Skhiri, 2001). Lip reading is something people do unconsciously and effortlessly. Listeners who watch a video wherein the speaker's audio output and visual lip movements match can accurately repeat what was said but 92 percent of listeners who receive conflicting audio output and visual lip movements repeat back incorrectly (McGurk & MacDonald, 1976). This phenomenon is called the McGurk effect. The addition of an image of a lecturer's face and lip movement to an auditory lesson might allow the listener to receive the message more accurately and to enhance learning of the lecture material. The image variable in the present study is used to

check the effect lip reading and dual channel coding has on learning.

Paivio's (1990) dual coding theory suggests that verbal and non-verbal processing are two distinctly specialized cognitive subsystems used in processing information. One subsystem processes images that Paivio calls "imagens", while the second subsystem processes what he calls "imogens" or verbal utterances. The underlying assumption of his dual coding theory is that all learners can process information through the two subsystems simultaneously without interfering with or blocking each other. Perhaps auditory (spoken) and visual (text) presentation of the same message could complement one another in this way.

Other research supports the view that dual coding of audiovisual and subtitled text works especially well for people who have a problem understanding audio input. Spanos and Smith (1990) reported that closed-captioned video allowed less proficient students to understand individual words from either an audio or video track whereas more proficient students were able to process language from both tracks. The less proficient students appeared to use either the audio or video track. Attempting to attend to both tracks might burden cognitive resources of the less proficient students to a point of being a distraction to the processing of the information. This supports their hypothesis that hearing impaired learners or learners with listening problems might benefit from having information processed through a visual medium. People who have no trouble understanding audio information might unconsciously use the medium that supports their individual learning style but people who do have trouble processing auditory information would have the opportunity to process the information through an alternative medium.

There is considerable time, effort, and expense involved in producing subtitled audiovisual lessons. This paper studies the effect of adding subtitles to an audiovisual taped lesson. The hypothesis was mainstream distance education learners who use subtitled audiovisual taped lessons will achieve at a higher level than those who do not.

Method

Participants

The participants were first year Algoma University College students who have successfully completed a short literacy-screening test. There were 16 participants per condition for a total of 64 participants. The literacy rankings were the basis for random assignment to one of four matched groups. The top four ranked participants were randomly assigned to one of the four condition groups. The next four highest ranked participants were randomly assigned to one of the four groups. The same random assignment procedure was applied to the remainder of the participants in groups of four, from the top rank downward.

Design and Procedure

A between- subjects 2 X 2 factorial design with an independent variable of image (present or not present) and an independent variable of text (present or not present) produced four conditions: audio only (blank screen), audio with an image of the lecturer presenting, audio with text, audio with text plus an image of the lecturer presenting.

Pretest: A multiple choice vocabulary pretest based on the "Word Power" vocabulary tests that are found in the Readers Digest magazine was manufactured.

The pretest consisted of 20 words with four definition choices per word.

Lecture presentation: A 6 minute taped lecture on velvet worms was produced. The lecturer's facial expressions and lip movements were observable, but there were no visual aids such as pictures, graphs, or props used in the lecture. Four versions of the taped lecture were produced. One had the audio lecture on a blank black television screen. The second taped lecture had audio with white printed subtitled text of what the lecturer was saying appearing on a black television screen. The third version had audio with an image of the lecturer presenting the lecture appearing on the television screen. The fourth version had audio with an image of the lecturer presenting the lecture and white printed subtitles of what the lecturer was saying appearing on the television screen.

A posttest consisting of 10 multiple-choice questions and a 10-point matching exercise based on the content of the lecture was manufactured and administered immediately after each lecture presentation.

Results

The mean posttest scores for each group are presented in graph 1. The four group means are displayed from highest to lowest: audio with text (mean = 14.5), audio only (mean = 13.75), audio with text and image (mean = 12.63), audio with image (mean = 11.9).
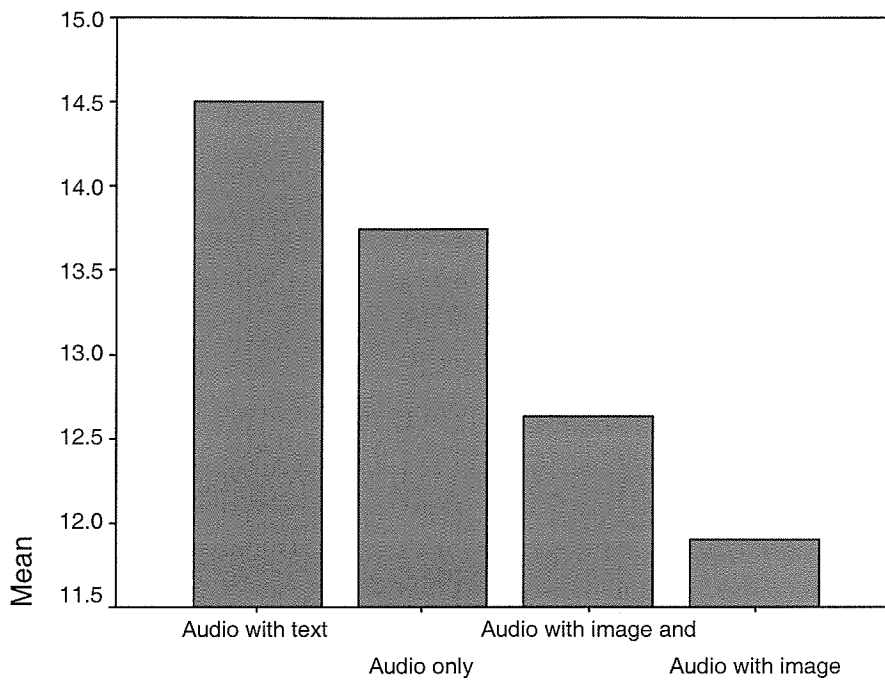
Graph 1

The data were analyzed by a factorial analysis of variance, F = [(3,60), ns]. No statistically significant effects or interactions were found.

Discussion

Although there was no statistically significant differences found between the groups, there were some interested results. The audio with text group had the highest group average followed by the audio only group. The audio with image and text group was third. The audio with image group had the lowest mean score. This is interesting because what might be considered ordinary videotape, without subtitles, is probably the most common form of audio video lecture material used in present distance education presentations.

The results were surprising, considering the highly significant differences Garza (1991) found in his research. Possibly the content of Garza's research (popular

television programming) or his sample population (students with high competence levels in the subject matter) could explain his statistically significant findings.

The results of the present student, although not significant, might be construed as to give weight to the argument that the addition of an image factor to an auditory lesson might tax the learner's cognitive processing. An attempt to replicate Garza's research might be one possible avenue for future research on the effects of learning from same language subtitled video.

References

Garza, T. J. (1991). Evaluating the use of captioned video materials in advanced foreign

language learning. *Foreign Language Annals 24*(3), 239-258.

McGurk, H. and MacDonald, J. (1976). *Hearing lips and seeing voices*. Nature, 264,

746-748.

Paivio, A. (1990). Mental Representations. New York: Oxford University Press.

Skhiri, M. (2001). Visual cues in speech perception. Retrieved November 19, 2003,

from Linkoping University, Swedish National Graduate School of Language

Technology Web site:

http:www.speech.kth.se/~rolf/gslt_papers/MustaphaSkiri.pdf

Spanos, G., & Smith, J. (1990). Closed captioned television for adult LEP literacy

learners. *Electronic Resources Information Center Digest*. Washington, DC:

National Center for ESL Literacy Education. (EDRS No. ED 321 623).