
Effects of Multimedia Courseware Subtitling on the Speaking Performance of College Students of French

ISABEL BORRÁS

*French Education Project
202 Peabody Hall
Louisiana State University
Baton Rouge, LA 70803
Email: nobile@marais.math.lsu.edu*

ROBERT C. LAFAYETTE

*French Education Project
202 Peabody Hall
Louisiana State University
Baton Rouge, LA 70803
Email: cilaja@lsuvm.sncc.lsu.edu*

OPPONENTS OF THE USE OF SUBTITLED video in foreign/second language teaching argue that the presence of subtitles is distracting and that they slow down the development of learners' listening abilities. Proponents of subtitles, on the other hand, contend that subtitles may help develop language proficiency by enabling learners to be conscious of language that they might not otherwise understand. Arguments advanced by both opponents and proponents of subtitles are largely intuitive and equally characterized by a lack of empirical evidence. This study aimed at providing some of that evidence by investigating the effects of subtitling during transactional task practice with multimedia courseware on oral communicative performance of fifth-semester college students of French. As used in this study, the term subtitling refers to whether the subjects can see and control, via interactive videodisc, a video segment with or without fully duplicating (literal), intralingual (in the target language) subtitles.¹ The study also investigated the effects of task level upon oral communicative performance and the interaction between subtitling and task level.

REVIEW OF RELATED RESEARCH

Most foreign/second language studies on the issue of subtitling have concentrated on the ef-

fects of subtitles on comprehension in reading or listening. None has investigated the effects of subtitles on speaking.

Price's study investigated whether exposure to subtitled video significantly improved or impaired video comprehension, as measured by a video post-test. The study also investigated which features of a student profile would influence performance on individual test items. Results from some five hundred students with more than twenty native language backgrounds indicated that "viewers, regardless of educational level or language background, benefited significantly from captioning, even with only one viewing" (p. 8).

In Vanderplank's study (42), fifteen European and eight Arabic learners of English, between high-intermediate and advanced level, watched nine hour-long sessions of BBC general output television programs with English subtitles. Subjects were asked to give various kinds of detailed feedback and to accomplish a number of language-oriented post-activities, such as acting out dialogues and scenes from situations seen on the programs. Results of the reports basically indicated that subjects found the subtitles beneficial to their language development, while results of the post-activities revealed that subjects had a high level of retention and recall of the language used in the programs. From these results, Vanderplank concluded that the use of subtitles may lead to the development of a "chunking ability" in both reading and listening, which in turn may release spare capacity for conscious language learning.

In the twelve-week study conducted by Neu-

man (7), 129 seventh- and eighth-grade students from Asian and Hispanic backgrounds were randomly assigned to one of four groups: 1) viewing closed-captioned television; 2) viewing television only; 3) reading and listening to the text; and 4) reading alone. Twice a week subjects either viewed or read segments of "3-2-1," a science series produced by the Children's Television Workshop. They were then tested on their ability to recognize and comprehend ninety target words. Students in the closed-captioned group scored higher than those in any other group.

Concerned with the generalizability of findings of second language (L2) research on subtitling to the "less commonly taught" foreign languages, Garza (11) conducted a comparative study using Russian and English (ESL) as target languages. Subjects in the study were forty third- and fourth-year students of Russian at two American universities and seventy high-intermediate/low-advanced ESL students at the same institutions. Video test materials for the experiment consisted of two sets of five Russian and American video segments, lasting sixteen and eighteen minutes respectively. Subjects in the Russian and American classes were randomly divided into two groups, so that one saw the test segments (with subtitles), and the other saw the control segments (without subtitles). Results of scores of all students on ten content-based comprehension checks indicated an overall gain in correct responses for subtitled versions of the ESL and particularly the Russian video segments. Similarly, results of random coded transcripts of the retelling showed a positive impact of subtitles on the subjects' language facility and a greater propensity to use the original lexicon of the video segments among subjects exposed to subtitles.

The above studies have shown the usefulness of subtitles for bridging the gap between the L2 learners' competence in reading and listening. None, however, has addressed the potential usefulness of subtitles for increasing learners' oral communicative performance. Contrary to Garza's assertion, the fact that subtitles may increase memorizability of language does not warrant "the student's ability to use that language in the proper context" (11: p. 245). For that to occur, students should be provided not only with subtitles but also with meaningful practice tasks supported by contextualized linguistic (grammatical and phonological) cues.

METHOD

Subjects. Forty-four students from the three sections of French 2154 and from one of the four sections of French 2155 at Louisiana State University (LSU) constituted the experimental sample. French 2154 (a reading course) and French 2155 (a conversation course) are both the fifth in a series of five courses that serve as an introduction to French and as a fulfillment of the foreign-language requirement for some majors in different Colleges of LSU. Both courses are the approximate equivalent of the end of college intermediate-high level study.

Design & Treatments. According to Salomon (33; 34), experiments on the instructional utility of media should investigate the ways in which a medium's coding elements and related learning tasks interact and affect learning.² Referring to a coding element such as subtitles, Vanderplank (42) suggests the investigation of their effects on productive skills in foreign/second language. As concerns the learning task, Krendl and Watkins contend that its manipulation can alter the learner's preconceptions about the difficulty of the medium and, therefore, increase the learner's mental effort and achievement. Based on these suggestions and assumptions, this study investigated both coding element and task variables by using a 2 × 2 factorial design. No control group was used in this experiment since the no-subtitles condition of the independent variable of primary interest, subtitling, also served as a control for no-treatment effects.

Subjects were randomly assigned to one of the treatments below based on the combination of the subtitling variable (subtitles *vs.* no-subtitles) with the oral task level variable (higher *vs.* lower):

- T1) subtitled video during oral task practice/lower-level task;
- T2) unsubtitled video during oral task practice/lower-level task;
- T3) subtitled video during oral task practice/higher-level task;
- T4) unsubtitled video during oral task practice/higher-level task.

Variables & Instrumentation. The first independent variable, *subtitling*, was a manipulated, active variable with two levels: 1) *subtitles*—subjects under treatments T1 and T3 used subtitled video segments to accomplish the required oral tasks; and 2) *no-subtitles*—subjects under treatments T2 and T4 used the same

video segments, but without subtitles, to accomplish the same oral tasks.

The second independent variable, *task level*, was also manipulated and active. It consisted of two oral transactional tasks—description and narration—each with two levels of difficulty, lower and higher. The levels were based on the amount and depth of information from the video segments required to accomplish the tasks successfully. *Lower-level tasks* were defined as tasks that required the subjects to provide a description or narration that included at least one spatial detail and three specific elements about the people or events they had seen in the video segments. *Higher-level tasks* asked the subjects to provide a description or narration that included various spatial details and at least five specific elements about the people or events they had seen in the video segments, which were identical to those seen by the lower-level task subjects.

Since the higher-level tasks posed more cognitive demands on the viewers than the lower-level tasks, we assumed that they would generate more complex talk. To accomplish a Description Unit, for instance, the higher-level task subjects, like their lower-level task counterparts, had to describe, based mainly on what they had seen in the video segment, the backdrop, physical appearance, and clothes of the main character. In addition, the higher-level task subjects had to report on the identity of the main character and make inferences about his/her personality, based mainly on what they had heard and on their own judgement.

All task directions were given in English to assure that what was tested was the subjects' ability to process and convey information in French and not the subjects' comprehension of the directions.

The two tasks, description and narration, were selected for two basic reasons. First, they illustrate the kinds of transactional tasks which according to Brown and Yule may allow learners to produce "long turns," an important element in the development of communicative abilities.³ Second, given the subjects' proficiency level, both tasks presumably could be accomplished, at least minimally, by the majority of the subjects.

The dependent variable, *oral communicative performance*, was defined as achievement in conveying information orally (describing people, objects, locations; narrating events, actions).⁴ This variable was measured by applying the instrument developed by Borrás (4), a semi-direct

test of speaking ability, to the oral samples produced by the subjects. The speaking test was designed such that the elicitation of spoken production was held constant, thus fulfilling the major requirement of the task-based approach to oral assessment used in this study. Four criteria were selected for measuring oral achievement at both global and component levels. At the global level, "effectiveness" was the major criterion by which subjects were to be evaluated. At the component level, "accuracy" was selected for the grammatical component, and "organization" and "fluency" for the discourse component.

The above criteria were identified and evaluated through the use of four six-point rating scales (see Appendix A) built on the following models: 1) Bartz's Amount of Communication and Accuracy scales were used to design the Effectiveness and Accuracy measures; 2) O'Malley et al.'s speech organization criteria were taken into account in building the Organization scale; 3) Emmett's Fluency scale provided the basis for the Fluency measure; and 4) Simon's communicative measurement procedure inspired the form of the evaluating instrument. The interrater reliability coefficients of the evaluating instrument were as follows: $r = .96$ (Overall Oral Performance); $r = .83$ (Effectiveness); $r = .95$ (Accuracy); $r = .94$ (Organization); and $r = .91$ (Fluency).

Materials. The courseware package entitled *Practicing Spoken French (PSF)* was used for the experimental task practice.⁵ The package, developed by Borrás (5) to meet the particular requirements of the experiment, integrates HyperCard and videodisc technologies and consists of two programs, Program 1 (P1) and Program 2 (P2), which allow for the individual practice of two video-based oral tasks—description and narration.⁶ Within each program, users can "customize" their ways of navigating through the units in terms of both length and pace.

Procedure & Data Collection. Subjects were assigned to the four treatment conditions resulting from the combinations of the subtitling and task level variables. Using the *PSF* package, subjects worked during two sessions on two optional units. Each unit was comprised of four working stages that focused on both the encoding of video-based information and the written and oral transmission of such information. The stages were as follows:

1) *Viewing:* Subjects watched a video segment at least two times. The segment was either sub-

titled or subtitled according to the particular subject's treatment;

2) *Checking comprehension input*: Subjects answered the questions of either "A game" (Description Unit) or "A quiz" (Narration Unit) about the information seen and heard in the video segment;

3) *Drafting one's own oral sample*: Subjects prepared for the production of an oral sample based on the segment they had seen. The preparation consisted of writing a draft of a description or a narration, following the higher/lower requirements of the task corresponding to the treatment to which subjects had been assigned. Throughout this stage, users were able to watch the video again or get linguistic help from the "Très utile" stack;

4) *Recording one's oral sample*: Subjects recorded the samples they had previously drafted. They were told that they could use their written draft as a reference, but not necessarily to read it aloud. Only five recordings were allowed, each for a maximum length of three minutes.

Following the tenets of cognitive psychology about meaningful learning (20; 43), stages one to three were aimed at ensuring that subjects were adequately prepared prior to undertaking the last and most important of the stages, speaking. We assumed that the activities of these three stages would help subjects improve their comprehension of the video segments and their memory of the people and events in those segments. It was equally assumed that the four stages would give subjects the opportunity to apply some of the cognitive strategies used by good second language communicators (26), namely, analyzing (analyzing expressions, translating), creating structure for input and output (taking notes, summarizing), and practicing (repeating, formally practicing with sounds and writing systems, recognizing and using formulas and patterns, recombining).

At the end of the second experimental session, subjects also filled out one of two forms, A or B, of a Likert questionnaire aimed at evaluating their attitudes toward the speaking practice with the *PSF* package.⁷ Questionnaire form A (see Appendix B) was administered to subjects in the subtitles condition (T1, T3) while questionnaire form B was administered to the remaining subjects (T2, T4). Since it was assumed that subjects were unaware that they had practiced under a variety of treatments, the two forms were written in a manner so as to maintain such unawareness. Statements one through three, identical for both questionnaires, ad-

ressed subjects' attitude toward their practice with multimedia courseware. Statements four through ten targeted subjects' opinion about particular aspects of the design of *Practicing Spoken French*: these statements were also identical in both questionnaires, with the exception of those asking for the subjects' opinion about subtitles (statements eight and nine, questionnaire form A) or videodisc features (statements eight and nine, questionnaire form B).

The oral samples were scored to provide a measure of the subjects' performance on the speaking practice tasks. For each sample, a composite score was obtained by adding the individual scores from the four six-point scales of the evaluating instrument (Effectiveness, Accuracy, Organization, and Fluency). Since "effectiveness" was considered the primary criterion of oral communicative performance, the points gained in the Effectiveness scale were weighed by three. Incorrect or inaccurate information did not count against the scores obtained on the different scales.

Experimental and descriptive data were gathered in four ways: 1) from the subjects' recorded oral samples; 2) from the measurements unobtrusively recorded by the computer; 3) from the researchers' observation of the subjects' oral sample drafts and transcripts as well as subjects' interaction with *PSF* and their reactions to the experimental site; and 4) from the two forms of the Likert questionnaire. Although experimental results were based only on the scores obtained in the second experimental session, scores of the first session were also taken into account to assess subjects' progress from one session to another, and to verify further the soundness of the experimental results.

Scores of oral samples recorded by the experimental subjects in the two experimental sessions were submitted to two-way factorial analyses of variance (ANOVA).⁸ The analyses were aimed at testing the main effects and interactions of the two independent variables, subtitling and task level. Tukey-Kramer post-hoc tests were computed to test the significance of mean differences for the independent variables main effects. Other two-way ANOVA analyses and ulterior Tukey-Kramer tests were run to assess the effects of subtitling and task type in both oral performance and comprehension questions ("A game"/"A quiz"). Pearson-r was used to investigate the reliability of the oral performance rating instrument, and the correlation between subjects' oral performance and attitude toward the computerized speaking practice.

The descriptive data provided by the simultaneous computer measurements were submitted to mean analysis for each of the treatment cells as well as for subjects across treatments (i.e., subtitles *vs.* no-subtitles, lower-level tasks *vs.* higher-level tasks). Subjects' attitude ratings were submitted to a one-way ANOVA to determine if they were affected by the subjects' assignment to the subtitles or the no-subtitles treatments.

MAIN FINDINGS

Analysis of the experimental data revealed statistically significant effects for subtitling, the independent variable of primary interest. Indeed, compared to the no-subtitles condition, the subtitles condition resulted in significantly higher overall Oral Performance scores in the two courseware programs; P1 [$F(1, 40) = 74.60, p < .001$] and P2 [$F(1, 40) = 68.41, p < .001$] (see Tables I and II).

The subtitles condition also resulted in four significantly higher subscores in both programs: Effectiveness [$F = 41.28, p < .001$ (P1); $F = 40.29, p < .001$ (P2)]; Accuracy [$F = 43.13, p < .001$ (P1); $F = 25.88, p < .001$ (P2)]; Organization [$F = 23.25, p < .001$ (P1); $F = 31.69, p < .001$ (P2)]; and Fluency [$F = 33.81, p < .001$ (P1); $F = 33.38, p < .001$ (P2)].

Results of Tukey-Kramer tests, illustrated in

Tables III and IV, demonstrated that in both programs the subtitles/lower-level task combination was statistically different at the $p < .05$ level from both the no-subtitles/lower-level task condition, and the no-subtitles/higher-level task condition at the $p < .05$ level. The combination subtitles/higher-level task was also significantly different at the $p < .05$ level from the no-subtitles/higher-level task condition.

Neither the effects of task level nor the interaction of this variable with the subtitles variable was significant. However, a consistent pattern of better performance of higher-level task subjects *vs.* lower-level task subjects was found in both programs. In both programs, as illustrated in Figures I and II, subjects in the subtitles/higher-level task cell also outperformed subjects in the other three cells.

In combination with an additional independent variable, task type, subtitles also appeared to have a significant effect upon oral communicative performance of the experimental subjects. Subtitles subjects did significantly better than no-subtitles subjects in the descriptive and the narrative tasks, both in P1 [$F(1, 40) = 69.41, p < .001$] and P2 [$F(1, 40) = 57.797, p < .001$]. In both programs the subtitles subjects who worked on the narrative tasks performed better than those who worked on the descriptive tasks. Post-hoc test analyses showed that, in both programs, the subtitles/descrip-

TABLE I
Summary ANOVA on Overall Oral Performance Scores (P1) for the Subtitling and Task Level Independent Variables

Source	df	Sum of Squares	Mean Squares	F-Value	P-Value
Subtitling	1	297.502	297.502	74.602***	.0001
Task Level	1	14.316	14.316	3.590	.0654
Sub. * T. L.	1	.041	.041	.010	.9195
Residual	40	159.514	3.988		

Dependent: Overall Oral Performance Scores (P1) *** $p < .001$

TABLE II
Summary ANOVA on Overall Oral Performance Scores (P2) for the Subtitling and Task Level Independent Variables

Source	df	Sum of Squares	Mean Squares	F-Value	P-Value
Subtitling	1	396.651	396.651	68.410***	.0001
Task Level	1	7.511	7.511	1.295	.2618
Sub. * T. L.	1	7.024	7.024	1.211	.2776
Residual	40	231.926	5.798		

Dependent: Overall Oral Performance Scores (P2) *** $p < .001$

FIGURE I
Interaction Bar Chart of Overall Performance Scores (P1) per Subtitling and Task Level Effects

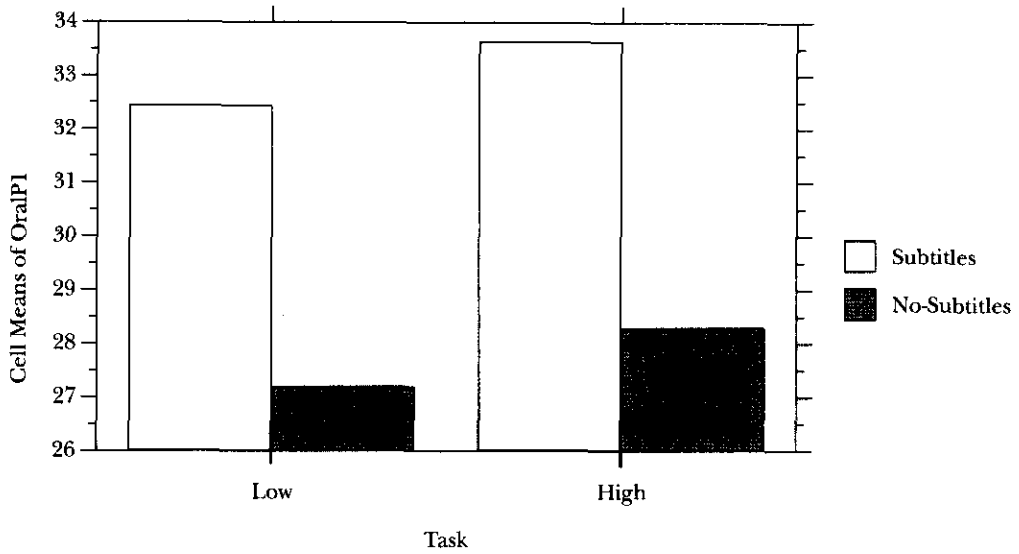


FIGURE II
Interaction Bar Chart of Overall Performance Scores (P2) per Subtitling and Task Level Effects

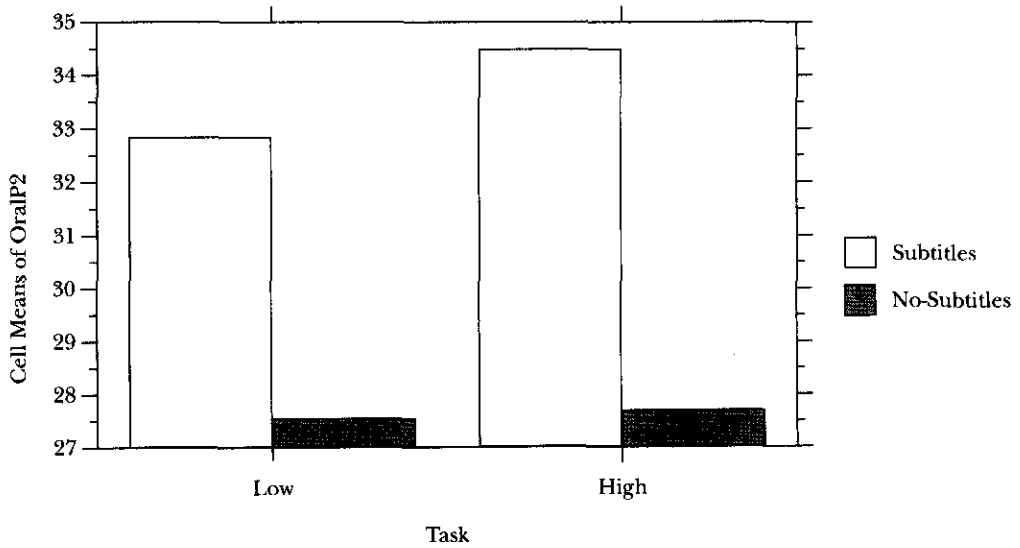


TABLE III
All Pairwise Comparisons of Means
(Tukey-Kramer's) on Overall Oral Performance
Scores (P1) for the Subtitling and Task Level
Independent Variables

A X B pairs	Mean Differences	Q _{statistic}
A ₁ B ₁ X A ₂ B ₁	5.15	8.58*
A ₁ B ₁ X A ₁ B ₂	-1.21	-2.05
A ₁ B ₁ X A ₂ B ₂	4.07	6.90*
A ₂ B ₁ X A ₁ B ₂	-6.36	-10.26
A ₂ B ₁ X A ₂ B ₂	-1.08	-1.74
A ₁ B ₂ X A ₂ B ₂	5.28	8.80*

* $p < .05$

Key: A₁ = Subtitles B₁ = Lower-level Task
A₂ = No-subtitles B₂ = Higher-level Task

TABLE IV
All Pairwise Comparisons of Means
(Tukey-Kramer's) on Overall Oral Performance
Scores (P2) for the Subtitling and Task Level
Independent Variables

A X B pairs	Mean Differences	Q _{statistic}
A ₁ B ₁ X A ₂ B ₁	5.22	7.15*
A ₁ B ₁ X A ₁ B ₂	-1.64	-2.31
A ₁ B ₁ X A ₂ B ₂	5.19	7.30*
A ₂ B ₁ X A ₁ B ₂	-6.86	-9.27
A ₂ B ₁ X A ₂ B ₂	-0.03	-0.04
A ₁ B ₂ X A ₂ B ₂	6.82	9.34*

* $p < .05$

Key: A₁ = Subtitles B₁ = Lower-level Task
A₂ = No-subtitles B₂ = Higher-level Task

tion condition was statistically different at the $p < .05$ level from the no-subtitles/description and the no-subtitles/narration conditions. The combination subtitles/narration was also significantly different from the combination no-subtitles/narration.

ADDITIONAL FINDINGS

An analysis of the data gathered from the measurements recorded by the computer, the observations of the researchers, and the attitude questionnaires provided additional insight into the experimental results of this study.

Computer Measurements. Computer data about subjects' total working time (see Figure III) in both programs indicated that subtitles/higher-level task subjects spent more time working on a program unit than subjects in the other three

cells. Data also indicated that the total working time of subjects in all cells decreased from P1 to P2, yet the overall Oral Performance scores of all cells, except T4, increased from the first to the second program.

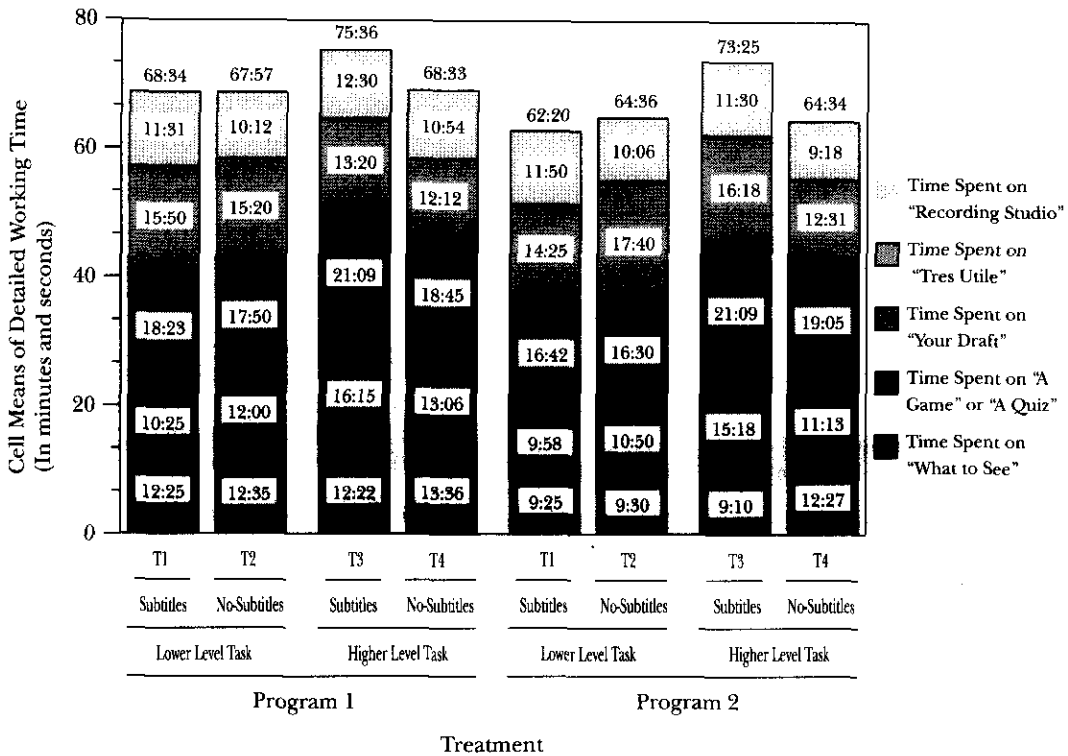
Computer data about subjects' activity-specific working times (see Figure III) provided the following information. First, contrary to what might be expected, the subtitles subjects spent less time in the first stage of the main navigating path of a program unit—initial viewing of a video segment—than the no-subtitles subjects. Second, subtitles subjects spent more time than no-subtitles subjects on the remaining stages of the main path—answering the comprehension questions "A game" or "A quiz," drafting the oral sample, and recording the oral sample. Third, the subtitles/higher-level task subjects spent more time in the optional navigating path, "*Très utile*," than their no-subtitles/higher-level task counterparts. All subjects, except the subtitles/lower-level task ones, stayed longer in "*Très utile*" in P2 than in P1.

Computer data about how frequently subjects entered "*Très utile*" indicated that the higher-level tasks subjects entered this optional path more often than did the lower-level task subjects; and among the former, the subtitles subjects entered this path more often than the no-subtitles subjects. Data also revealed that the lower-level task subjects recorded their samples more often than did the higher-level task subjects.

Researchers' Observations. The four kinds of observations recorded by the authors during the two experimental sessions (interaction with PSF, reactions to the experimental site, oral sample drafts and oral sample transcripts) led to interesting findings. First, the interaction of subjects with the PSF courseware was qualitatively and quantitatively greater than expected. In spite of their lack of familiarity with HyperCard, subjects needed little help to proceed with their units and invariably followed the stages and the directions of those units.

Second, the reactions of subjects to the experimental site were quite positive. According to their comments, the majority of the students preferred the computer laboratory where the experiment took place to the language laboratory, where they usually view video excerpts and have CALL practice sessions for their French classes. They seemed to enjoy sharing the site with students who were using computers for disciplines other than French. Being in the computer laboratory with students working on

FIGURE III
Interaction Bar Chart of Total and Activity-Specific Working Times per Treatment and Program



a variety of subject matters, as opposed to being in the language laboratory surrounded by students doing similar exercises, gave them a feeling of independence and creativity.

Third, observation of the subjects' oral sample drafts revealed that: 1) oral sample drafts that had been typed on the computer (fifty-four percent in P1, and fifty-eight percent in P2) were generally better organized and had fewer spelling errors than drafts which had been written by hand; 2) reuse of lexical items seen in the "Très utile" stack was very high, particularly among the subtitles/higher-level task subjects, who made extensive use of the connectors (adverbs, conjunctions) included in the "How to Narrate" section of that stack; and 3) reuse of difficult words such as numbers or

proper names heard in the video segments was, as could be expected, higher among the subtitles than the no-subtitles subjects.

Attitude Questionnaires. Data from the Likert questionnaires revealed a highly positive attitude among experimental subjects toward speaking practice with multimedia courseware and, in particular, toward the *Practicing Spoken French* package. This attitude turned out to be significantly higher for the subtitles subjects than for the no-subtitles ones. The one-way (subtitles vs. no-subtitles) ANOVA performed on attitude ratings revealed a significant subtitles effect [$F(1, 42) = 4.249, p < .05$]. Little to low correlation was found, however, between the attitude and the oral performance of subjects in the four treatments.

STUDY FINDINGS & CALL ISSUES

The descriptive data collected during this study also contributed to the existing research knowledge on the following CALL issues: 1) the effects of type of control on performance; 2) the informational value of time on task; 3) the effects of word processing on quality of written production quality; and 4) the relationship between attitude toward and performance with multimedia courseware.

Type of Control. The design of the courseware package used in this study, *Practicing Spoken French*, used a combined learner/program control formula that, on the one hand, provided learners with control over the pace of the lesson, the type of learning module, the video and the audio segments, the textual information resources, and whether they elected to receive feedback on the video comprehension questions. On the other hand, the program controlled the depth of the task, the access to the subtitles, and the number of recordings.

This formula seems to have worked as effectively as the formulae featured by the materials created for previous CALL (30) and CAI (14; 2) studies—combined learner-program control and learner-control with advisement, respectively. Computer data about the total and detailed amount of time that the subjects spent on a *PSF* unit indicated that the combination learner/program control was successful in keeping subjects “at task” during the allotted time, thereby ensuring that their exposure to the units was sufficient to detect any experimental effect.⁹

Computer information about the high frequency with which subjects recorded their oral samples or entered the “*Très utile*” stack seems to indicate the active involvement of subjects in two actions that were left to their control—recording and entering an optional path.

Time on Task. Research on this issue has shown that time on task may be an indicator of the cost-effectiveness of CALL programs (25) and, more importantly, of the learning strategies used by subjects during CALL practice (27). Corroborating those findings, the data collected in this study of the time spent by subjects in the different stages of a *PSF* unit seems to indicate that: 1) *PSF* was cost-effective; and 2) *PSF* motivated subjects to make use of the cognitive strategies that had been incorporated in the package’s design.

PSF was effective in reducing learning time across the two experimental sessions. Computer

records indicated that all subjects spent less time in P2 than in P1, yet their oral performance scores and subscores were higher in the second program than in the first.

Subjects made use of the strategies embedded in *PSF* such as analyzing, creating structure for input and output, and, especially, practicing. Data about the total and activity-specific working times of both programs indicated that all subjects worked consistently in both the main and the optional paths of the *PSF* unit for at least the minimal time the authors had estimated they would spend in those paths.

Word Processing. Foreign/second language research has shown the positive impact of word processing on the quality of written production and on attitudes toward the writing process (37; 23). Results from this study underscore those findings. Indeed, researchers’ observations of the subjects’ oral sample drafts indicated that subjects who used word processing were more organized and accurate than those who wrote by hand. Also, subjects who typed on the computer seemed to enjoy the process of revising their drafts.

Attitudes. Data from subjects’ comments and answers to the Likert questionnaire also contributes to the existing body of knowledge about attitudes toward CALL. First, subjects’ attitudes toward the speaking practice with *PSF* were very positive but correlated poorly with subjects’ oral performance. This finding supports those reported by Roblyer, Castine, and King’s meta-analysis of attitudes toward CAI. Second, contrary to Ahmad, Corbett, Rogers, and Sussex’s findings, subjects in this study with little experience in using computers reacted as favorably toward the computerized speaking practice as did those with more experience.

DISCUSSION & CONCLUSIONS

The results from this study, summarized in Table V, support and advance the knowledge generated by foreign/second language research on subtitles. Indeed, these results seem to indicate that far from being detrimental, fully duplicating intralingual subtitles have potential value in helping the learner to not only better comprehend authentic linguistic input but also to produce comprehensible communicative output.

The statistically significant difference found in this study in favor of the subtitles condition for higher oral communicative performance strongly suggests that when learning from

TABLE V
Significant F-Ratios for the Subtitling Variable Yielded by the ANOVA Analyses

Independent Variable(s)	Dependent Variable	F-Value
<i>Program 1</i>		
Subtitling/Task Level	Overall Oral Performance	74.602***
Subtitling/Task Level	Effectiveness	41.284***
Subtitling/Task Level	Accuracy	43.132***
Subtitling/Task Level	Organization	29.245***
Subtitling/Task Level	Fluency	33.808***
Subtitling/Task Type	Overall Oral Performance	69.411***
<i>Program 2</i>		
Subtitling/Task Level	Overall Oral Performance	68.410***
Subtitling/Task Level	Effectiveness	40.290***
Subtitling/Task Level	Accuracy	25.882***
Subtitling/Task Level	Organization	31.690***
Subtitling/Task Level	Fluency	33.384***
Subtitling/Task Type	Overall Oral Performance	57.797***
Subtitling	Attitudes Ratings	4.249*

*** $p < .001$ * $p < .05$

"authentic video" in a multimedia environment, having the opportunity to see and control subtitles, as opposed to not having that opportunity, results in both better comprehension and subsequent better productive use of the foreign language.

The significant effects of subtitles support Salomon's theory (33) about how media symbol systems and coding elements affect cognition and learning. Within the context of that theory and in light of the experimental results, it can be concluded that: 1) video with fully duplicating intralingual subtitles may help the foreign/second language learner associate the aural and written forms of words more easily and quickly than video without subtitles; 2) empowering the foreign/second language learner with the control of subtitles, through such interactive technologies as interactive videodisc (IVD),¹⁰ may contribute to the internalization of this coding element and the short-cutting of the skill (simultaneous listening and reading) that it facilitates; and 3) the simultaneous activation of listening and reading skills by the subtitles may require a certain level of mastery of those skills particularly if subtitles are to be relevant to the specific learning task of speaking. Therefore, intermediate/advanced students may benefit more from subtitles than beginning students.

Additional experimental and descriptive data reinforce the above conclusions or lead to additional ones. The fact that subtitles subjects spent about the same or less time watching the video segments at the beginning of the program

than their no-subtitles counterparts indicated that they did not pay much attention to the subtitles the first time they saw them. This inattention, which was probably related to the subjects' unfamiliarity with this coding element—Americans rarely watch subtitled foreign movies—suggests that practice using native language subtitles in informal learning situations may improve students' attitudes toward and the effective use of target language subtitles in academic foreign/second language situations.¹¹ The fact that the attitudes of subtitles subjects were significantly better than those of no-subtitles subjects leads to the conclusion that subtitles may increase learners' Perceived Self-Efficacy (PSE) (32) and promote better attitudes toward mediated language learning. The subtitles subjects in this study may have perceived themselves as being quite capable of obtaining information from a medium that provided them with the transcript of the aural source.

Although the effects of the secondary independent variable, task level, were insignificant, a pattern of better performance for the higher-level task subjects than for lower-level task subjects was found in both programs. This pattern, which underscores results of previous research on instructional media about the effects of perception of task difficulty on mental effort and performance (16; 35), might have been significant if the sample size had been greater.

Although the effects of the interaction of subtitling and task level were also insignificant, the subtitles/higher-level task treatment (T3)

turned out to be the most effective. T3 subjects obtained greater scores in the oral communicative performance measure than subjects in the other three cells. Moreover, T3 subjects stayed longer "at task" and produced longer oral samples than subjects in the other treatments. Finally, in comparison with other subjects, T3 subjects seemed to retain and to reuse better the more complex vocabulary seen on "*Très utile*" (connectors) or read in the video segments (numbers and proper nouns). These observations indicate that the subtitles may be all the more effective if they are high-level task oriented.

IMPLICATIONS & RECOMMENDATIONS

This study has shown that allowing fifth-semester college students of French the possibility of seeing and controlling subtitles may increase their performance on video-based oral communicative practice tasks with multimedia courseware. Since the study is probably the first of its kind, this finding needs further empirical assessment. The finding, however, should move the profession to consider the four potential implications of subtitles in foreign/language teaching and learning. First, with the addition of the self-control feature, this study has put subtitles under a new pedagogical light. Self-controlled subtitles may empower learners to "adjust" the redundancy of the aural and visual channels of video programs according to the learning requirements or language mastery, thus extending the potential effectiveness of the subtitles to the early stages of language learning and to either visually- or aurally-oriented learners. Second, against arguments that contend that early introduction of writing may be negative (22), and in favor of positions that argue for the promotion of formal accuracy of learner output (18; 29), subtitles could make both the comprehension of "authentic" input and the production of "accurate" oral/written output less "painful." Third, developing learners' "taste" for subtitles would increase learners' interest in maintaining and/or increasing their knowledge of the foreign language outside the academic setting (e.g., through the use of domestic foreign broadcast videotext). And fourth, subtitles could be an effective way to promote reading literacy, particularly among learners of the so-called developed countries where communication technology has made individuals very receptive to audiovisual information—sometimes at the risk

of leading them to superficiality—and almost always at the expense of depriving them of the enlightening powers of the written word.

A number of issues warrant further investigation.

1) The study should be replicated with a larger sample in order to create sufficient statistical power to discern the probably significant effects of the task-level variable.

2) Since the study has evaluated oral performance based on only two types of transactional tasks and four video segments, replication of the study with different tasks and segments would also be advisable.

3) The study has proven that giving learners control over the pace of the subtitles may be an effective way to reduce channel layer density and improve the quantity and quality of learners' oral intake and production. Further research on the issue should investigate whether the additional possibility of inserting and removing the subtitles may increase such benefits.

4) The long-term effects of speaking practice under subtitles *vs.* no-subtitles treatments is also a promising area for future research. Subjects participated in this experiment for less than 180 minutes, yet they obviously experienced the learning benefits of subtitles. The investigation of the long term effects of subtitles during speaking practice with multimedia courseware may provide additional useful information about the learners' internalization of this new "mental tool" and its incorporation in the learners' cognitive repertoire.

5) Additional research should be conducted to investigate how the font type or the color coding of subtitles might serve as a reliable guide for the identification of the various speech sources of a video segment. While font type was not truly investigated in this study due to equipment limitations, text color coding has been tested in computer programs for teaching grammar (21).

6) The investigation of the video segment length variable may also be an interesting avenue of research. In this study, the length of the subtitled video segments (two to four minutes) was based on recommendations from previous works (17; 12) about optimal length of unsubtitled video to maintain L2 learner's attention. It would be interesting to determine how the increase of this optimal length with subtitled video affects learners' attention span.

If it is to be shown that it is "the *application* of the technology of education rather than the

provision of technology in education" (38: p. 21) that affects learning performance, then the above recommendations have to be put into practice.¹² The implementation of studies similar to the one described here will not be without difficulties either materially (development of the experimental materials and implementation of the experimental procedure may be very time consuming), financially (significant investment in hardware, authoring tools, and interfaces may be necessary), or logistically (solicitation of sponsoring institutions for the development of the courseware and its empirical assessment may require management and expertise). Nevertheless, significant efforts should be made to carry out these studies. The challenge of such efforts is to identify the appropriate speaking practice to be presented in the appropriate form to the appropriate students. The opportunity is to improve computerized language learning and education in general.

NOTES

¹ Following Jung's terminology, "intralingual" subtitles differ from "interlingual" subtitles, or translation subtitles, in that they are unilingual renditions of the soundtrack. Both interlingual and intralingual subtitles can be "partially" or "fully duplicating," referring to how much of the spoken dialogue/monologue or voice-over commentary is translated or transcribed.

² Salomon defines media as "our cultural apparatus for selecting, gathering, storing, and conveying knowledge in representational form" (33: p. 3). According to him, any medium is made up of many symbol systems, which in turn are made up of various coding elements. For example, in a medium such as video, one important symbol system is "display." The coding elements of the "display" symbol system would include, among others, graphics, images (size, color), and script information (subtitles: font, size, color).

³ Brown and Yule delineate a variety of "tasks" that students can perform to demonstrate their speaking ability. The tasks reflect a range of speech modes used in talking about static (e.g., describing an object or photograph), dynamic (e.g., giving an eye-witness account), and abstract relationships (e.g., justifying a course of action).

⁴ As used in this study, the term "communicative" does not necessarily imply the idea of oral interchange with a listener but rather the transmission of information of a given amount and depth that can be comprehended by that listener. Also, as used in this study, the term "performance" relates to the acquisition of specific linguistic features that are associated with a particular course of instruction. It contrasts

with the term "proficiency," which refers to an acquired competence in the language regardless of how or where the language was acquired.

⁵ *Practicing Spoken French* was run in four workstations housed in the computer laboratory of the College of Education at LSU. Each of the workstations included the following hardware components: 1) a Macintosh IIsi computer with 5mb RAM and 80mb Internal Hard Disk; 2) an AppleColor High-Resolution RGB monitor (13"); 3) a Pioneer LD-V2200 videodisc player; 4) an Amdex Color 300T TV monitor; and 5) a Personal LaserWriter LS printer.

⁶ The visual input for *Practicing Spoken French* was provided by four video segments selected from *France Panorama* (10), a quarterly video magazine containing contemporary "authentic" materials from France and designed for use in school and university French second-language programs. The experimental segments were selected following the evaluation of sixteen video segments by four instructors who were teaching French 2154 at LSU. The selection was carried out through the use of an evaluating instrument designed by Borrás (4) drawing on the works of Garza (12) and Grimes. The instrument consisted of four rating scales that measured the segments in terms of three "quality" criteria (situational appropriateness, inherent interest value, audio/visual correlation) and one "difficulty" criterion (grammatical and lexical complexity). The overall quality and difficulty mean ratings for the segments retained for the first experimental session were 4.08 and 2.87 respectively (one on the quality scales was "very poor" and five was "excellent"; one on the difficulty scale was "very easy" and five was "very difficult"). The quality and difficulty mean ratings for the segments of the second session were 4.12 and 3.00 respectively. After selection, the four video segments were transferred to a SMPTE 1" tape, overlaid with subtitles, and pressed onto a videodisc.

⁷ Likert's criterion of internal consistency (19) was computed on the responses of the subjects of the pilot study that preceded the experiment to one of the two forms of the questionnaire. The criterion showed the questionnaire's high ability to discriminate between positive and negative reactions to speaking practice with the *Practicing Spoken French* package.

⁸ The statistical analyses of this study were carried out using the *SuperANOVA* software package, and the *SPSS X Data Analysis System*, Release 3.0.

⁹ Evidence exists that under a learner-control condition, students tend to spend less time on the lesson than subjects in a learner-control with advisement condition (14).

¹⁰ Within the interactionist perspective of Second Language Acquisition (SLA), the potential of IVD for language learning has been extensively highlighted. Doughty, for instance, observes that "... the interaction between learner and the auditory and visual environment created in the IVD context may be hypothesized as facilitative to the second language acquisition process" (p. 3).

¹¹ In an article about non-native speakers' use of teletext subtitles in English, Vanderplank (41) comments on the "good" language ability and the positive attitude of the general population of the Nordic countries and Holland toward viewing foreign language programs, which are invariably accompanied by interlingual subtitles.

¹² Unfortunately, as Spencer points out, provision of technology "is frequently given priority when non-educational agencies are involved, ill-informed decisionmakers being seduced by the superficial appeal of new hardware approaches" (p. 21).

BIBLIOGRAPHY

1. Ahmad, Khurshid, Greville Corbett, Margaret Rogers & Roland Sussex. *Computers, Language Learning and Language Teaching*. New York: Cambridge Univ. Press, 1985.
2. Arnone, Marilyn P. & Barbara L. Grabowski. "Effects on Children's Achievement and Curiosity of Variation in Learner Control Over an Interactive Video Lesson." *Educational Technology Research and Development* 40 (1992): 15-27.
3. Bartz, Walter H. *Testing Oral Communication in the Foreign Language Classroom*. Arlington, VA: Center for Applied Linguistics, 1979.
4. Borrás, Isabel. "Effects of Subtitled Video during Transactional Task Practice on Oral Communicative Performance of Fifth Semester College Students of French Learning with Multimedia Courseware." Unpub. Diss., Louisiana State Univ., 1993.
5. ———. "Practicing Spoken French (PSF): Multimedia Courseware for Improving Oral Communicative Skills." *CALICO '93: Assessment*. Ed. Frank L. Borchardt & Eleanor M. T. Johnson. Durham, NC: Duke Univ. Press, 1993: 16-20.
6. Brown, Gillian & George Yule. *Teaching the Spoken Language: An Approach Based on the Analysis of Conversational English*. Cambridge: Cambridge Univ. Press, 1983.
7. "Closed-Captioned TV Is Effective Teaching Tool." *ACTFL Newsletter* 3 (1991): 6-7.
8. Doughty, Catherine. "Theoretical Motivations for IVD Software Research and Development." *Interactive Videodisc: The "Why" and the "How."* CALICO Monograph Series 2. Ed. Michael D. Bush, Alice Slaton, Miguel Verano & Martha E. Slayden. Provo, UT: Brigham Young Univ. Press, 1991: 1-15.
9. Emmett, Allan. "The Associated Examining Board's Test in English for Educational Purposes (TEEP)." *L'Evaluation de la "Performance" en Langue Seconde*. Ed. P. C. Hauptman, R. Leblanc & M. B. Wesche. Ottawa: Univ. of Ottawa Press, 1985: 131-51.
10. *France Panorama* [video magazine]. New York: Eagle Multimedia Services, 1991.
11. Garza, Thomas J. "Evaluating the Use of Captioned Video Materials in Advanced Foreign Language Learning." *Foreign Language Annals* 24 (1991): 239-58.
12. ———. "Foreign Language Teaching and Video: Providing a Context for Communicative Competence." Unpub. Thesis, Harvard Graduate School of Education, 1986.
13. Grimes, Tom. "Audio-video Correspondence and Its Role in Attention and Memory." *Educational Technology Research and Development* 38 (1990): 15-25.
14. Johansen, Keith J. & Robert D. Tennyson. "Effect of Adaptive Advisement on Perception in Learner-controlled, Computer-based Instruction Using a Rule-Learning Task." *Educational Communication and Technology Journal* 31 (1983): 226-36.
15. Jung, Udo O. H. "The Challenge of Broadcast Videotex to Applied Linguistics." *International Review of Applied Linguistics in Language Teaching (IRAL)* 28 (1990): 201-20.
16. Krendl, Kathy A. & Bruce Watkins. "Understanding Television: An Exploratory Inquiry into the Reconstruction of Narrative Content." *Educational Communication and Technology Journal* 31 (1983): 201-12.
17. Lavery, Mike. *Active Viewing Plus*. Oxford: Modern English Publications, 1984.
18. Leech, Geoffrey N. *Semantics*. Harmondsworth, UK: Penguin, 1974.
19. Likert, Rensis. "Techniques for the Measurement of Attitudes." *Archives of Psychology* 140 (1932): 5-55.
20. McDonough, Steven H. *Psychology in Foreign Language Teaching*. London: Allen & Unwin, 1981.
21. Meara, Peter. "Lexical Skills and CALL." *Computers in English Language Teaching*. Ed. Christopher Brumfit. Oxford: Pergamon, 1985: 83-89.
22. Muller, Daniel. "The Effect upon Pronunciation and Intonation of Early Exposure to the Written Word." *Modern Language Journal* 49 (1965): 411-13.
23. Neu, Joyce & Robin Scarcella. "Word Processing in the ESL Writing Classroom: A Survey of Students' Attitudes." *Computer-Assisted Language Learning and Testing: Research Issues and Practice*. Ed. Patricia Dunkel. New York: Newbury House, 1991: 169-87.
24. O'Malley, J. Michael, Anna U. Chamot, Gloria Stewner-Manzanares, Rocco P. Russo & Lisa Küpper. "Learning Strategy Applications with Students of English as a Second Language." *TESOL Quarterly* 19 (1985): 557-84.
25. Orlansky, Jesse. "Effectiveness of CAI: A Different Finding." *Electronic Learning* 3 (1983): 58-60.
26. Oxford, Rebecca L. *Language Learning Strategies: What Every Teacher Should Know*. New York: Newbury House, 1990.
27. Pederson, Kathleen M. "The Effects of Passage Availability During Adjunct Questioning in

- Computer-assisted Reading Practice on Recall Measures of Reading Comprehension in Intermediate College French." Unpub. Diss., Ohio State Univ., 1985.
28. Price, Karen. "Closed-Captioned TV: An Untapped Resource." *MATSOL Newsletter* 12 (1983): 7-8.
 29. Rivers, Wilga M. "Reflections on Language Learning and Teaching: An Interview with Wilga Rivers." *English Teaching Forum* 39 (1991): 2-5.
 30. Robinson, Gail L. "Effective Feedback Strategies in CALL: Learning Theory and Empirical Research." *Computer-Assisted Language Learning and Testing: Research Issues and Practice*. Ed. Patricia Dunkel. New York: Newbury House, 1991: 155-67.
 31. Roblyer, M. D., W. H. Castine & F. J. King. *Assessing the Impact of Computer-Based Instruction: A Review of Recent Research*. Binghamton, NY: Haworth, 1988.
 32. Salomon, Gavriel. "The Different Investment in Mental Effort in Learning from Different Sources." *Educational Psychologist* 18 (1983): 42-51.
 33. ———. *Interaction of Media, Cognition and Learning*. Washington: Jossey-Bass, 1979.
 34. ———. "What Is Learned and How It Is Taught: The Interaction between Media, Message, Task, and Learner." *Media and Symbols: The Forms of Expression, Communication and Education*. Ed. D. R. Olson. National Society for the Study of Education 73. Chicago: Univ. of Chicago Press, 1974: 383-406.
 35. ——— & Tamar Leigh. "Predispositions about Learning from Print and Television." *Journal of Communication* 34 (1984): 119-35.
 36. Simon, Charlann S. *Evaluating Communicative Competence: Functional Pragmatic Procedure*. Rev. ed. Tucson: Communication Skill Builders, 1986.
 37. Sommers, Elisabeth A. "Integrating Composing and Computing." *Writing On-line: Using Computers in the Teaching of Writing*. Ed. James L. Collins & Elisabeth A. Sommers. Upper Montclair: Boynton/Cook, 1985: 3-10.
 38. Spencer, Ken. "Modes, Media and Methods: The Search for Effectiveness." *British Journal of Educational Technology* 22 (1991): 12-22.
 39. *SPSS X Data Analysis System*, Release 3.0 [software]. Chicago: SPSS, Inc., 1988.
 40. *SuperANOVA* [software]. Berkeley, CA: Abacus Concepts, 1991.
 41. Vanderplank, Robert. "Non-Native Speaker Viewers' Use of Teletext Subtitles in English: Some Exploratory Studies." *New Technology in Language Learning*. Ed. G. Davies & M. Husey. Frankfurt: Lang, 1992: 89-102.
 42. ———. "The Value of Teletext Subtitles in Language Learning." *ELT Journal* 42 (1988): 272-81.
 43. Wittrock, Merlin C. "Learning and the Brain." *The Brain and Psychology*. Ed. Merlin C. Wittrock. New York: Academic, 1980: 371-403.

 APPENDIX A

 Oral Communicative Performance: Rating Scales

EFFECTIVENESS SCALE (adapted from Bartz, 1979)

General definition: Amount of relevant information conveyed by the speaker and understood by the listener.

Definition of each level of the scale:

1. Virtually no relevant information was conveyed by the subject.
2. Very little relevant information was conveyed by the subject.
3. Some relevant information was conveyed by the subject.
4. A fair amount of relevant information was conveyed by the subject.
5. Most relevant information was conveyed by the subject.
6. All relevant information was conveyed by the subject.

Note: Read carefully the "Directions for the subject" of all tasks to know what exactly "relevant information" means on each one of these tasks.

ACCURACY SCALE (adapted from Bartz, 1979)

General definition: The grammatical correctness of subject's utterances.

Definition of each level of the scale:

1. No utterances rendered correctly.
2. Structure of very few utterances rendered correctly.
3. Some utterances rendered correctly, but many structural problems remain.
4. Many correct utterances, but some problems remain with structures.
5. Most utterances rendered correctly; only minor problems with structures.
6. All utterances rendered correctly.

Note: When applying this scale to the narrative samples, the evaluator should take into account that subjects were required to use the present tense throughout their speeches.

ORGANIZATION SCALE

General definition: The overall coherence and cohesion of the subject's speech (O'Malley et al., 1985).

Definition of each level of the scale:

1. No descriptive detail or narrative event is presented according to an order. No helpful inter-sentential connectors.
2. Very few details or events are presented according to an order. Few sentences are properly interconnected.
3. Some details or events are presented following an order. Some sentences are properly interconnected.
4. Many details or events are sequentially presented.

A reasonable range of proper intersentential connectors is used.

5. Most details or events presented according to an order. Most of the sentences are well interconnected.
6. All details or events are rendered according to an order. All sentences are properly interconnected.

Note: Read carefully the information on the "Très utile" stack to learn about the "order" and the connectors suggested to the subjects.

FLUENCY SCALE (adapted from Emmett, 1985)

General definition: The overall smoothness, continuity, and naturalness of the subject's speech as opposed to pauses for rephrasing sentences, groping for words, and so forth.

Definition of each level of the scale:

1. Utterances so halting and fragmentary that communication is virtually impossible.
2. Utterances very slow, uneven, and often incomplete.
3. Utterances fairly slow, hesitant, and uneven.
4. Utterances produced at a reasonable speed though with occasional hesitations.
5. Utterances quite fast and fairly effortless.
6. Utterances produced with a native speaker's speed and ease.

APPENDIX B

Attitude Questionnaire Form A

The information you provide in this questionnaire will be of great value for the designer of the multimedia package *Practicing Spoken French*. Please do not sign the form.

Directions: Rate on a scale of 1 to 5 the following questions. Please circle only one number for each category.

Rating values: 1 (strongly disagree); 2 (disagree); 3 (no opinion); 4 (agree); 5 (strongly agree).

1. Overall, I found the French oral practice sessions with multimedia courseware to be a worthwhile learning experience.

1 2 3 4 5

2. I would like to have had the opportunity to take more practice sessions like these ones.

1 2 3 4 5

3. More practice sessions similar to the ones I had would increase my confidence to speak to other people in French.

1 2 3 4 5

4. This type of speaking practice, with the "free navigation" through the modules and the video materials, enabled me to get through the required oral tasks, description or narration, by myself.

1 2 3 4 5

5. The characters and events portrayed on the video segments were interesting.

1 2 3 4 5

6. The questions and the feedback of the trivial pursuit and quizzes helped me to better understand and remember the content of the video segments.

1 2 3 4 5

7. The information on the "Très utile" stack provided me with the vocabulary and grammar needed to prepare the descriptive or the narrative tasks.

1 2 3 4 5

8. The subtitles increased my comprehension of what was said on the video segments, and therefore, my ability to express the content of these segments.

1 2 3 4 5

9. The two styles of the subtitles (slanting, upright) helped me to distinguish between what was said by people seen and people heard on the video images.

1 2 3 4 5

10. The time and the number of tries allowed to record the description or the narration were sufficient.

1 2 3 4 5

Copyright of Modern Language Journal is the property of Wiley-Blackwell and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.