1993

Effects of Subtitled Video During Transactional Task Practice on Oral Communicative Performance of Fifth Semester College Students of French Learning With Multimedia Courseware.

Isabel Borras
Louisiana State University and Agricultural & Mechanical College

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_disstheses

Recommended Citation

This Dissertation is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Historical Dissertations and Theses by an authorized administrator of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.
INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

University Microfilms International
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
313/761-4700 800/521-0600

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Effects of subtitled video during transactional task practice on oral communicative performance of fifth semester college students of French learning with multimedia courseware

Borrás, Isabel, Ph.D.
The Louisiana State University and Agricultural and Mechanical Col., 1993

Copyright ©1993 by Borrás, Isabel. All rights reserved.
EFFECTS OF SUBTITLED VIDEO DURING TRANSACTIONAL TASK PRACTICE ON ORAL COMMUNICATIVE PERFORMANCE OF FIFTH SEMESTER COLLEGE STUDENTS OF FRENCH LEARNING WITH MULTIMEDIA COURSEWARE

A Dissertation

Submitted to the Graduate Faculty of the Louisiana State University and Agricultural and Mechanical College in partial fulfillment of the requirements for the degree of Doctor of Philosophy in The Department of Curriculum & Instruction

by Isabel Borrás
B.A., University of Valencia, Spain, 1976
M.A., University of Valencia, Spain, 1985
D.E.S.S., University of Poitiers, France, 1987
May 1993
To my parents, Juan Bautista and Carmen, and to my sister Mª Carmen, whose love and encouragement made this project possible.
ACKNOWLEDGMENTS

I would like to thank the many people whose support has made possible this pursuit.

I am deeply grateful to my adviser, Robert, C. Lafayette. His knowledge, careful counsel and encouragement played an important role in this project and throughout my doctoral studies.

I wish to acknowledge the members of my committee. I am indebted to Carl Blyth for his interest and assistance in the early and final stages of the research. And I thank Susan MacGregor, Arnulfo Ramirez, and William E. Doll, Jr. for their reading of the manuscript and their most helpful input.

I wish to thank the people who contributed to the creation of the multimedia package "Practicing Spoken French" used in this project. In particular, I wish to thank Alexander Perlis for his efficiency in programming the package and subtitling the video segments, and his assistance in helping to solve the technical problems that emerged during the implementation of the project. I thank Robert C. Lafayette, Annette Tamuly, Denise Egéa, and Carl Blyth, who "provided" their voices and proofread the text; Salem Chenafi, Daniel Earle and Gary Dauphin who took or scanned the pictures; Fred Glickman who gave the permission to use the copyrighted video segments; and the companies that mastered and pressed the videodisc.
My heartfelt thanks also go to those who provided their time and experience throughout the research project. I thank Robert C. Lafayette and Carl Blyth for serving as validators of the video segments, and Catherine Jolivet for her assistance in the validation and application of the evaluating instrument.

I am grateful to the students of French 2154 and 2155, and to their instructors, Jefferson Humphries, Hélène Moreau, Frédéric Pallez, and Adélaïde Russo, who encouraged them to participate. I acknowledge the College of Education for allowing me to conduct the experiment at its computer laboratory.

I wish also thank the following people for their encouragement or assistance throughout the project: Augusto Nobile, Erena Messina, Beth Harris, Joyce Gregory, Gail Lungaro, Udo Jung, Anil Pauloxc, and Brett Brock.

Finally, my especial thanks go to my family. I thank my father, my mother, and my sister for their steadfast and loving support, without which this project would never have come to fruition.
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Tables</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1 Mean Ratings of Preselected Video Segments on Quality and Difficulty Criteria</td>
<td>124</td>
</tr>
<tr>
<td>5.1 Summary ANOVA on Overall Oral Performance Scores (P1) for the Subtitling and Task Level Independent Variables</td>
<td>137</td>
</tr>
<tr>
<td>5.2 Summary ANOVA on Overall Oral Performance Scores (P2) for the Subtitling and Task Level Independent Variables</td>
<td>137</td>
</tr>
<tr>
<td>5.3 Summary ANOVA on Effectiveness Scores (P1) for the Subtitling and Task Level Independent Variables</td>
<td>138</td>
</tr>
<tr>
<td>5.4 Summary ANOVA on Effectiveness Scores (P2) for the Subtitling and Task Level Independent Variables</td>
<td>138</td>
</tr>
<tr>
<td>5.5 Summary ANOVA on Accuracy Scores (P1) for the Subtitling and Task Level Independent Variables</td>
<td>139</td>
</tr>
<tr>
<td>5.6 Summary ANOVA on Accuracy Scores (P2) for the Subtitling and Task Level Independent Variables</td>
<td>139</td>
</tr>
<tr>
<td>5.7 Summary ANOVA on Organization Scores (P1) for the Subtitling and Task Level Independent Variables</td>
<td>140</td>
</tr>
<tr>
<td>5.8 Summary ANOVA on Organization Scores (P2) for the Subtitling and Task Level Independent Variables</td>
<td>140</td>
</tr>
<tr>
<td>5.9 Summary ANOVA on Fluency Scores (P1) for the Subtitling and Task Level Independent Variables</td>
<td>141</td>
</tr>
<tr>
<td>5.10 Summary ANOVA on Fluency Scores (P2) for the Subtitling and Task Level Independent Variables</td>
<td>141</td>
</tr>
</tbody>
</table>
5.11 All Pairwise Comparisons of Means (Tukey-Kramer's) on Overall Oral Performance Scores (P1) for the Subtitling and Task Level Independent Variables .......................... 145

5.12 All Pairwise Comparisons of Means (Tukey-Kramer's) on Overall Oral Performance Scores (P2) for the Subtitling and Task Level Independent Variables .......................................... 145

5.13 Summary ANOVA on Overall Oral Performance Scores (P1) for the Subtitling and Task Type Independent Variables .................. 147

5.14 Summary ANOVA on Overall Oral Performance Scores (P2) for the Subtitling and Task Type Independent Variables .................. 147

5.15 All Pairwise Comparisons of Means (Tukey-Kramer's) on Overall Oral Performance Scores (P1) for the Subtitling and Task Type Independent Variables .......................... 149

5.16 All Pairwise Comparisons of Means (Tukey-Kramer's) on Overall Oral Performance Scores (P2) for the Subtitling and Task Type Independent Variables .......................... 149

5.17 Summary ANOVA on Questions ("A Game"/ "A Quiz") Scores (P1) for the Subtitling and Task Type Independent Variables .................. 159

5.18 Summary ANOVA on Questions ("A Game"/ "A Quiz") Scores (P2) for the Subtitling and Task Type Independent Variables .................. 159

5.19 All Pairwise Comparisons of Means (Tukey-Kramer's) on Questions ("A Game"/ "A Quiz") Scores (P1) for the Subtitling and Task Type Independent Variables .................. 162

5.20 All Pairwise Comparisons of Means (Tukey-Kramer's) on Questions ("A Game"/ "A Quiz") Scores (P2) for the Subtitling and Task Type Independent Variables .................. 162

5.21 Summary ANOVA on Attitude Ratings for the Subtitling Variable ......................................................... 167

vii
5.22 Mean Statement Scores of Subtitles Cells (T1, T3) on Attitude Questionnaire Form A ........................................................... 168
5.23 Mean Statement Scores of No-Subtitles Cells (T2, T4) on Attitude Questionnaire Form B ........................................................... 169
6.1 Significant F-Ratios for the Subtitling Variable Yielded by the ANOVA Analysis .......................................................... 183
6.2 Significant Q-Statistics Yielded by the Tukey-Kramer Tests Performed on the ANOVA Results .............................................. 184
N.1 Means and Standard Deviations of Overall Oral Performance Scores (P1) as a Function of Subtitling and Task Type ........... 267
N.2 Means and Standard Deviations of Effectiveness Scores (P1) as a Function of Subtitling and Task Type .................. 267
N.3 Means and Standard Deviations of Accuracy Scores (P1) as a Function of Subtitling and Task Type .................. 268
N.4 Means and Standard Deviations of Organization Scores (P1) as a Function of Subtitling and Task Type ................ 268
N.5 Means and Standard Deviations of Fluency Scores (P1) as a Function of Subtitling and Task Type .................. 269
O.1 Means and Standard Deviations of Overall Oral Performance Scores (P2) as a Function of Subtitling and Task Type ........... 270
O.2 Means and Standard Deviations of Effectiveness Scores (P2) as a Function of Subtitling and Task Type .................. 270
O.3 Means and Standard Deviations of Accuracy Scores (P2) as a Function of Subtitling and Task Type .................. 271
O.4 Means and Standard Deviations of Organization Scores (P2) as a Function of Subtitling and Task Type .................. 271
O.5 Means and Standard Deviations of Fluency Scores (P2) as a Function of Subtitling and Task Type .................. 272
P.1 Means and Standard Deviations of Overall Oral Performance Scores (P1) as a Function of Subtitling and Task Type .......... 273

P.2 Means and Standard Deviations of Overall Oral Performance Scores (P2) as a Function of Subtitling and Task Type .......... 273

Q.1 Means and Standard Deviations of Questions ("A Game","A Quiz") Scores (P1) as a Function of Subtitling and Task Type ........................................ 274

Q.2 Means and Standard Deviations of Questions ("A Game","A Quiz") Scores (P2) as a Function of Subtitling and Task Type ........................................ 274
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figures</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 2 X 2 Factorial Design</td>
<td>78</td>
</tr>
<tr>
<td>3.2 Oral Communicative Competence Construct</td>
<td>82</td>
</tr>
<tr>
<td>4.1 &quot;Practicing Spoken French&quot;: Menu-Driven Design</td>
<td>101</td>
</tr>
<tr>
<td>4.2 to 4.15 PSF screens</td>
<td>105-118</td>
</tr>
<tr>
<td>4.16 Flowchart Program 1. Description Unit</td>
<td>120</td>
</tr>
<tr>
<td>5.1 Interaction Plot of Overall Oral Performance Scores (P1) per Subtitling and Task Level Effects</td>
<td>144</td>
</tr>
<tr>
<td>5.2 Interaction Plot of Overall Oral Performance Scores (P2) per Subtitling and Task Level Effects</td>
<td>144</td>
</tr>
<tr>
<td>5.3 Interaction Bar Chart of Overall Performance Scores (P1) per Subtitling and Task Type Effects</td>
<td>148</td>
</tr>
<tr>
<td>5.4 Interaction Bar Chart of Overall Oral Performance Scores (P2) per Subtitling and Task Type Effects</td>
<td>148</td>
</tr>
<tr>
<td>5.5 Interaction Bar Chart of Total and Detailed Working Time per Treatment and Program</td>
<td>153</td>
</tr>
<tr>
<td>5.6 Interaction Bar Chart of Length Oral Sample per Treatment and Program</td>
<td>157</td>
</tr>
<tr>
<td>5.7 Interaction Bar Chart of Questions (&quot;A Game&quot;/&quot;A Quiz&quot;) Scores (P1) per Subtitling and Task Type Effects</td>
<td>161</td>
</tr>
<tr>
<td>5.8 Interaction Bar Chart of Questions (&quot;A Game&quot;/&quot;A Quiz&quot;) Scores (P2) per Subtitling and Task Type Effects</td>
<td>161</td>
</tr>
</tbody>
</table>
5.9 Bar Chart of Overall Reactions of Subtitles Cells (T1, T3) to Attitude Questionnaire Form A ....................................................... 171

5.10 Bar Chart of Overall Reactions of No-Subtitles Cells (T2, T4) to Attitude Questionnaire Form B ....................................................... 172

M.1 Interaction Bar Chart of Overall Oral Performance per Treatment and Program ................................................................. 262

M.2 Interaction Bar Chart of Oral Performance ("Effectiveness") per Treatment and Program ......................................................... 263

M.3 Interaction Bar Chart of Oral Performance ("Accuracy") per Treatment and Program ................................................................. 264

M.4 Interaction Bar Chart of Oral Performance ("Organization") per Treatment and Program ................................................................. 265

M.5 Interaction Bar Chart of Oral Performance ("Fluency") per Treatment and Program ................................................................. 266
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>x</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>xviii</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>1 INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>The Problem</td>
<td>1</td>
</tr>
<tr>
<td>Perspective on the Problem</td>
<td>2</td>
</tr>
<tr>
<td>Some Machine-Aided Solutions</td>
<td>2</td>
</tr>
<tr>
<td>A Working Alternative</td>
<td>4</td>
</tr>
<tr>
<td>Purpose of the Study</td>
<td>6</td>
</tr>
<tr>
<td>Rationale for the Design of the Study</td>
<td>9</td>
</tr>
<tr>
<td>Coding Element</td>
<td>11</td>
</tr>
<tr>
<td>Theoretical Base</td>
<td>11</td>
</tr>
<tr>
<td>Empirical Base</td>
<td>14</td>
</tr>
<tr>
<td>Learning Task</td>
<td>15</td>
</tr>
<tr>
<td>Task Level</td>
<td>15</td>
</tr>
<tr>
<td>Task Type</td>
<td>16</td>
</tr>
<tr>
<td>Definition of Terms</td>
<td>17</td>
</tr>
<tr>
<td>Assumptions and Limitations</td>
<td>21</td>
</tr>
<tr>
<td>Overview of the Dissertation</td>
<td>23</td>
</tr>
<tr>
<td>2 REVIEW OF RELATED LITERATURE</td>
<td>24</td>
</tr>
<tr>
<td>Introduction</td>
<td>24</td>
</tr>
<tr>
<td>Subtitling</td>
<td>27</td>
</tr>
</tbody>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
REFERENCES ................................................................. 212

APPENDICES ................................................................. 231
A Invitation Letter to Experimental Subjects .................. 232
B Oral Performance: Evaluation Form ......................... 233
C Subjects’ Oral Samples: Eight Transcripts ............... 243
D Attitude Questionnaire Form A ................................. 247
E Attitude Questionnaire Form B ................................. 249
F "Practicing Spoken French": Instructions to the User .... 251
G Courseware Development: Staff Hours ................... 253
H PSF Documentation .................................................. 255
I Preselected Video Segments ..................................... 256
J Instrument for Rating the Quality and Difficulty of the Experimental Video Segments ...................... 258
K Selected Video Segments: Mean Ratings and Length 260
L Video Segments Production: Who did What? ............ 261
M Interaction Bar Charts of Oral Performance Criteria per Treatment and Program ......................... 262
N Means and Standard Deviations of Overall Oral Performance Scores and Subscores (P1) as a Function of Subtitling and Task Level ........................................... 267
O Means and Standard Deviations of Overall Oral Performance Scores and Subscores (P2) as a Function of Subtitling and Task Level ........................................... 270

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
ABSTRACT

The purpose of this study was to investigate the effectiveness of subtitled video during transactional task practice on oral communicative performance of fifth semester college students of French learning with multimedia courseware.

Task practice was provided by the multimedia package, "Practicing Spoken French" (PSF) specifically designed by the researcher for this study. The package, which integrates HyperCard and videodisc technologies, allows for individually customized practice of two video-based oral tasks, description and narration.

Drawing on Salomon's theories about the effects of media coding elements and perception of task on cognition and learning, the study used a 2 x 2 factorial design with two levels of subtitling (subtitles, no-subtitles), and two levels of oral transactional tasks (higher, lower).

Forty-four fifth semester college students of French were randomly assigned to one of four experimental treatments -- subtitles/lower-level task, no-subtitles/lower-level task, subtitles/higher-level task, no-subtitles/higher-level task. Students under each treatment were required to participate in two experimental sessions to complete the four stages of the two practice tasks of their choice -- watching a video segment, answering video-related
questions, drafting a description or a narration following the video information, and recording an up to 3-minute oral sample, based on what they had drafted.

The dependent variable, oral communicative performance, was assessed by applying four six-point rating scales (Effectiveness, Accuracy, Organization, and Fluency) to the subjects' oral samples. Subjects in the subtitles treatments scored significantly higher on the oral performance measure and had significantly better attitudes than subjects in the no-subtitles treatments. The effects of task level and the interaction of this variable with subtitling were insignificant. However, higher-level task subjects outperformed their lower-level task counterparts in the two experimental sessions. Also in both sessions, subtitles/higher-level task subjects obtained the highest scores on the oral performance measure.

Insights on CALL issues such as locus of control, time on task, learning environment, word processing, and attitudes were offered, as well as conclusions and implications for speaking assessment, CALL design, and future research.
CHAPTER 1
INTRODUCTION

The Problem.

Oral communication is probably one of the most difficult aspects of foreign language teaching and learning. It is not an easy task to individually monitor students' oral practice, or to objectively evaluate students' oral production. In written production, for example, students can get on by themselves, without disturbing the rest of the class, at their own speed. In oral production, however, students need to speak individually. When a student speaks, he or she makes a noise that will disturb other students unless they are saying the same thing at the same time or unless they are listening to what he or she is saying. The possible ways of coping with this logistical drawback seem to be limited. First, students may give individual short responses to the teacher's questions. Second, students can practice individually with the help of technological equipment. And third, in rare privileged environments, students can engage in small-group "conversation" in the foreign language.
As concerns the evaluation of the oral production, it is not less problematic. Many well-established tests do not even have an oral component, since grammatical accuracy and vocabulary can be assessed quite adequately, it seems, in the written mode.

**Perspective on the Problem**

**Some Machine-Aided Solutions**

With respect to the above mentioned option of using technological equipment for individual oral practice, the language laboratory has been the most common but not the best solution. Based on notions of "correct responses," either with respect to pronunciation, grammar, or both, language laboratory practice may enable learners to improve their ability to produce short structured responses in familiar dialogue slots. What it obviously cannot prepare learners to do is to produce extended responses, since it gives them no opportunity to produce "long turns" which seem essential to the development of communicative abilities (Brown & Yule, 1983a).

The language laboratory and the related "audiolingual" and "audiovisual" methods were at their peak in the 1960s. But in the 1970s, as questions were raised about the theories underlying the new methods of teaching, doubts also began to be expressed about the merits of the
language laboratory for developing listening and speaking skills in a second/foreign language.

In the mid-1980s, just as the unsubstantiated panegyrics for the effectiveness of the language laboratory dampened the enthusiasm of second language teachers, unjustified claims about the effectiveness of computers have engendered the skepticism of some specialists about the value of Computer-Assisted Language Learning (CALL). Ariew (1987), for instance, in his discussion about the integration of CALL and video in the foreign language curriculum, denies the ability of computers to provide speaking practice when he says, "At this time, speaking can only be addressed by teachers making use of their experience and the examples, hints, and suggestions pedagogical guides and teachers' editions of textbooks typically provide." (p. 51). There are three reasons for this skepticism: first, traditional tests of communicative performance exert little control over the students' oral production; second, due to technological limitations, oral performance assessment cannot be done properly by the computer; and third, computerized speaking practice required, until very recently, the interfacing of tape recorders or sound digitizers (e.g., Farallon MacRecorder, 1987), which made the recording and playback of students' voices a lengthy and costly process.
More optimistic CALL specialists have, however, suggested that computer programs have potential for simulating conversation among pairs or groups of learners which in turn may enhance the development of second language communicative competence. However, initial research into the actual talk generated by the computer has shown the discourse to be limited in amount and complexity. Windeatt (1986), Piper (1986), Mydlarski (1987), and Abraham and Liou (1991) investigated the quantity and quality of "talk" produced by small groups of learners working in different text-based CALL programs (English Second Language/non-English Second Language). Somewhat discouraging, the results of these studies indicated that the actual realizations of students tended to be short and often involved repetitions of screen text or another group member's utterances. The CALL programs used in the studies did not create opportunities for learners to use the type of conversational language in which they could extend themselves linguistically.

A Working Alternative

Fortunately, the skepticism expressed by theorists and the meager research results are not sufficient arguments to dismiss the potential effectiveness of computers to improve oral communicative skills. With respect to the reasons behind theorists' skepticism, only the assessment problem which in itself seems to impede computerized speaking interaction
is not solved yet. The problems of oral ability assessment and computerized speaking practice could now respectively be solved through the use of task-based tests of communicative performance, and the use of built-in recording facilities available on the newest computers.

As concerns the research, two possible reasons for the lackluster results of the studies mentioned earlier could be the nature of the stimuli (text-based materials) and the type of courseware (market-available CALL programs) used to generate talk in such studies. Underscoring these reasons, Piper (1986) notes that the programs used in her study were chosen because of their availability rather than their appropriate quality. And she adds that those programs afforded little or no opportunity to practice "long turns" which she thinks, agreeing with Brown and Yule (1983a), to be of great value for increasing communicative ability.

A working alternative would, therefore, be to provide computerized oral practice based on different stimuli (authentic video) and different materials (purposely designed multimedia courseware). First, authentic video would provide better opportunities for learners to understand and imitate the target language presented in a real-like context. Second, purposely designed multimedia courseware would ensure the adequacy of the learning practice with the learning objective: orienting the learner more effectively toward the desired length and type of speech production.
As concerns the video stimuli, the adoption of a coding element, such as fully duplicating intralingual subtitles, that is, on-screen verbatim of the language spoken in a video segment, would probably affect oral output. Indeed, this coding element, which has already proved its value for improving reading and listening comprehension from videotaped materials (Price 1983; Garza, 1991), would be even of greater interest when used for speaking purposes. Giving learners the possibility of seeing and "controlling" the printed version of the narrative of a video, would increase learners' opportunities to "model" their oral productions meaningfully.

With respect to the multimedia courseware, it would be now possible to create a package which, combining the features of videodisc and HyperCard technologies, could very well motivate students not only "to try harder and learn more" (Underwood 1989: p. 19), but also to explore new ways of thinking and learning.

Purpose of the Study

The primary aim of this study was to investigate the effects of subtitled video during transactional task practice upon oral communicative performance of college students of French learning with a specifically designed multimedia package, "Practicing Spoken French" (PSF).
The main question of the study was the following:

- Does subtitled video vs. unsubtitled video during transactional task practice have any effect on oral communicative performance of fifth semester college students of French?

In order to evaluate effectively the likely effects of the primary independent variable, subtitling, the experiment also took into account another independent variable: the task level. The higher- or lower-level of a task was a function of the amount and depth of the information from a video segment required to accomplish such task.

Two other questions emerged from the combination of these two treatment independent variables:

- Do lower- vs. higher-level oral transactional tasks have any effect, not attributable to subtitles, upon oral communicative performance of fifth semester college students of French?
- Is there any interaction between the two independent variables, subtitling and task level?

In addition to the above primary research questions, a fourth one arose from the combination of the subtitles and the task type variables:

- Does subtitled video vs. unsubtitled video during descriptive vs. narrative task practice have any effect upon oral
communicative performance of fifth semester college students of French?

The primary and the additional questions of the study were selected to provide information about the most essential effects of the coding element, fully duplicating intralingual subtitles.

It was decided to study the effects of subtitles because of their potential ramifications. As previously mentioned, there is research evidence indicating that subtitles increase comprehension of the linguistic content of video materials on intermediate/advanced foreign language learners. However, nothing is known about the effects of subtitles when they can be "controlled" by learners working in a multimedia environment.

Similarly, it was decided to investigate the effects of tasks requiring differing levels of information processing since there is proof that perceptions of the difficulty level of a learning task may influence the "cognitive capacity usage" (Britton, Muth, and Glynn, 1986), or the amount of "mental effort" (Salomon, 1983) invested in that task.

Finally, a secondary aim of the study was to provide further insights on some basic research issues involving the effectiveness of CALL. These include:

- the effects of locus of control\(^{(3)}\) on performance;
- the informational value of time on task;
- the effects of learning environment on attitudes;
- the effects of word processing on written production quality;
- the relationship between attitude toward and performance with CALL.

**Rationale for the Design of the Study**

In the past two decades, media experts and foreign language educators have become increasingly aware of the fruitless nature of what have been coined "broad methodologies comparisons." Clark and Salomon (1983), stress that media do not affect learning in and of themselves, and that general media comparisons do not yield useful information. They exhort that studies pertaining to the overall instructional impact of media not be conducted. In the same way, Pederson (1987) states that,

Comparative research that endeavors to identify a superior medium or a superior method is doomed to provide murky and usually uninterpretable results because the essential variables that might make a difference in a given language-learning context cannot be isolated, manipulated, and examined adequately. (p. 101).

Generally speaking, appropriate research in Computer Assisted Instruction (CAI) has lagged far behind materials development. Although an impressive number of CAI applications have developed (Conklin, 1987; Morariu, 1988; Megarry, 1988; Jonassen, 1988a, 1988b, 1989), research on
educational computing has been characterized, as Green and Bigum (1990) observe, "... by a reductionistic epistemology which equates thinking with data." (p. 370). This is particularly true in the area of Computer Assisted Language Learning (CALL) where interesting materials have also lately appeared (Ashworth, D. & Stelovsky, J., 1989; Underwood, 1989; Chun & Brandl, 1992) but where there is still little information about how to use the computer capabilities to enhance lesson delivery (Meredith, 1983; Underwood, 1984).

To date, most CALL research has been either comparative (Teicher, 1985) or evaluative (Simonsen, 1985), with results somewhat poor either because of the impossibility of replicating the experimental conditions, in the first case, or the rapid obsolescence of software evaluations, in the second one.

As an alternative to the comparative or evaluative models, CALL basic research has suggested that it is the design of computer software intended to cause adjustments in cognitive processing, not the medium used to deliver instruction, that stands the best chance of affecting learning outcomes. A few CALL studies have been conducted (Pederson, 1985; Robinson, Underwood, Rivers, Hernández, Rudesill, & Enseñat 1985) that illustrate that learning benefits can be gained from taking into account
variables such as the learner, the learning task, and the coding elements of
the medium rather than the medium itself.

Contributing to this type of research, and drawing on Salomon's
(1979) model of mediated instruction, this study will investigate how a
particular type of medium coding element and learning task affect the
learning of a second language skill, speaking.

**Coding Element**

**Theoretical Base**

In his discussion about the instructional capabilities of hypermedia,
Jonassen (1988b) observes that hypertext "will find its most effective use as
a learning tool that reflects what [sic] learner knows...a mind tool, as it were,
and not a delivery vehicle" (p. 33). Jonassen's observation echoes
Salomon's (1974) assertion that when the educational effectiveness of
media is at issue, it is the media's "symbol system rather than the
technology of transmission that is crucial for instruction" (p. 385).

The parallelism between these two assertions, made in different
decades and referring to different media underscores the "relevance" of
Salomon's model (1979) of mediated instruction, and suggests its
application to the design of this study.

Salomon defines media as "our cultural apparatus for selecting,
gathering, storing, and conveying knowledge in representational form"
(1979, p. 3). Any medium, is made up of many symbol systems. For instance, in a medium such as video, one important symbol system is "display." Symbol systems in turn are made up of various subcategories or coding elements. The coding elements of the "display" symbol system would include, among others, color, graphics, font, size and type, and information format (scriptural/visual).

Salomon insists that no coding element can be attributed solely to one medium. Each medium, however, does "translate" coding elements in ways that result in some differences from their use by other media. For instance, with interactive videotext (IVD), a coding element such as "subtitles" can offer a broader range of expressive and communicational possibilities than with video only -- since IVD images can be instantly and randomly accessed. The possibilities are even greater when IVD becomes a component of a multimedia environment and the subtitles can be controlled from the computer keyboard and explored to fulfill meaning-enhancing tasks.

According to Salomon's theoretical propositions, symbol systems and coding elements vary with respect to: 1) the amount of mental translation required for the extraction and processing of knowledge; 2) the amount of mental skills required for that purpose, and 3) the mental skills
they cultivate. Let us look at these propositions and the way they apply to subtitles, the coding element in this study.

1. Symbol systems vary as to the cognitive systems they address, and given a particular content, person, and task, the information they carry requires different amounts of mental recoding and elaboration. As Salomon emphasizes, "the better the correspondence between the way information is presented and the way it can be mentally represented, the less recoding is needed and the easier is the communication." (1979, p. 86).

Therefore, verbatim target language subtitles may help learners to associate the aural and written forms of words more easily and quickly.

2. Symbol systems call upon different sets of mental skills for the extraction and processing of the coded information. And, as Salomon adds, "The effectiveness of employing a symbol system for instructional purposes depends on the extent to which the specifically selected coding elements activate in the learner mental skills that are sufficiently mastered by them and are relevant to the requirements of the task." (1979, p. 110).

Consequently, if subtitles are to be relevant to the required task, speaking, a certain level of mastery of the two skills which they are supposed to activate, reading and listening, will be required. Intermediate students would probably benefit more from subtitles than beginning students.
3. Symbol systems and coding elements may affect cognition in three ways: "activation" of skills that transform external codes into internal ones; "short-cutting" of skills by overtly providing the end result of mental transformations that the learner should have employed; and overt skill "supplantation" by means of transformational codes. According to Salomon, "Skill cultivation by activation is a typical procedure of evoking mental processes and reinforcing them, but short-circuiting and supplantation could turn a communicational code into a tool of thought" (1979, p. 137).

Hence, the learner's interaction with subtitles would account for the internalization of this coding element and the short-cutting of the skill, simultaneous listening and reading, that it facilitates.

**Empirical Base**

Adding to the theoretical grounds which support the choice of subtitles as the coding element in this study, four empirical reasons provide further support for that choice. First, there is sound research evidence (Omaggio, 1979; Mueller, 1980) that visuals in and of themselves do not necessarily improve comprehension in second language learning. Subsequently, to enhance video images with scriptural cues such as subtitles may lead to a better processing and understanding of the information channeled by such images. Second, studies in audio-video channel correspondence also support the hypothesis of greater
comprehension for subtitled video materials. Indeed, if it is true, as Grimes (1990) observes, that "when two information channels are correspondent, the narration 'drives' the processing of the visuals" (p. 24), it follows that subtitles could "fuel" such a driving. Third, as research indicates (Koran, Snow & McDonald, 1971), the written presentation of the narrative accompanying complex visual images may be particularly beneficial for those subjects which have difficulty in processing the images' verbal input on their own. Four, subtitles should include the verbatim of a video or motion picture since it seems that only exact duplicates on screen of what is being said in the narration lead to unimpeded processing of the intended message (Reese, 1982).

Learning Task

Task Level

Activation of specific skills depends not only on the coding elements but also on the tasks to be performed. The learner's task orientation, or the perception of what is supposed to be learned, determines the effect of a given coding element. As Salomon (1979) puts it, "Task requirements, whether imposed or self-selected, determine what kinds of information are to be extracted, and this choice determines in turn what kinds of coding elements within the message are to be addressed." (p. 108).
Salomon makes an interesting distinction between the cognitive effects and the effectiveness of coding systems when he says: "Whereas [cognitive] effects reflect the interaction between a symbol system and one's mastery of the needed mental skills, effectiveness reflects the interaction between the activated mental skills and the skill-requirement of the task" (1979, p. 110). Consequently, he argues that for effective instructional communication to take place, a match needs to be established between the cognitive demands of the task, the skills that are required by the codes of the message, and the learner's level of mastery of these skills.

Taking into account Salomon's viewpoint, the design of this study harmonizes the level of information processing required by the task (higher/lower) with the skill required by the subtitles (simultaneous listening and reading), and the subjects' expected level of command of this skill (high-intermediate).

**Task Type**

Two types of transactional tasks, description and narration, have been chosen for the experimental speaking practice of this study. The choice of these tasks is based on the following reasons: first, the tasks exemplify Brown and Yule's (1983a) categories of activities suitable for the task-based elicitation of spoken accounts of static (description) and dynamic (narration) relationships; second, the tasks allow the subjects to
"talk at length," thereby making possible the surfacing of subtle incompetent communicative features which standardized tests fail to show (Brown & Yule, 1983a; Simon, 1986); third, the tasks can be assessed through discourse-analysis techniques suggested by researchers such as Brown and Yule (1983b), van Dijk and Kintsch (1983), and Abraham and Liou (1991). Fourth, taping the informative language function described by Holliday (1973), the tasks may yield information about the subjects' abilities to 1) perceive, interpret, and comment on the stimuli presented in a video segment, and 2) selectively feature and subordinate details. Abilities which are considered to be a predominant characteristic of effective communicators (Loban, 1976; Simon, 1986).

### Definition of Terms

**ACTFL:** American Council on the Teaching of Foreign Languages.

**"A game"/*"A quiz":** Comprehension tests of the video segments used during the practice tasks. Included in the "Description" units of P1 and P2, "A game" is made up of five true/false or multiple-choice questions. A part of the "Narration" units of both programs, "A quiz" consists of five fill-in-the-blank, relating, and ordering questions.

**CALL:** Computer Assisted Language Learning.
<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAV disc:</td>
<td>Laser videodisc recorded at a Constant Angle Velocity (as opposed to a CLV disc recorded at a Constant Linear Velocity).</td>
</tr>
<tr>
<td>HyperCard:</td>
<td>The best known of the first generation of computer languages supporting creation of highly interactive multimedia programs. It consists of a collection of screens, or cards, which can be used for whatever purpose, and a collection of tools (drawing, painting, text layout, and &quot;clip art&quot;) based on two Macintosh graphics and word-processing programs (MacPaint and MacWrite).</td>
</tr>
<tr>
<td>HyperCard card:</td>
<td>Central object category on the HyperCard hierarchical structure. A screen, or piece, of textual and/or graphic information.</td>
</tr>
<tr>
<td>HyperCard background:</td>
<td>Object category which can be shared by various cards.</td>
</tr>
<tr>
<td>HyperCard stack:</td>
<td>Sequential accumulation of cards and backgrounds.</td>
</tr>
<tr>
<td>HyperCard script:</td>
<td>Sequence of instructions that tells objects how to respond. Every &quot;script,&quot; called &quot;program&quot; in other languages, must be attached to some object, either card or background.</td>
</tr>
<tr>
<td>HyperCard button:</td>
<td>Object sensitive to some user actions such as pointing and clicking on the mouse to such button.</td>
</tr>
<tr>
<td>hypermedia:</td>
<td>Nonlinear and interactive methods of moving through a body of information available in a variety of formats --e.g., text, graphics, audio, video.</td>
</tr>
<tr>
<td>hypertext:</td>
<td>Coined by Nelson in the 1960s, the term Hypertext basically describes a non-sequential mode of reading and writing that links different nodes of a text. The term also describes the</td>
</tr>
</tbody>
</table>
electronic representation of text that takes advantage of the random access capabilities of computers to overcome the strictly sequential medium of printed paper.

initial video processing time: The amount of time spent watching a video segment before answering the related video comprehension questions, "A game" or "A quiz."

interactive videodisc (IVD): Instructional media system which combines the power of the microcomputer with the image and the audio storage capabilities of the optical laser disc.

medium: Any form of technology of transmission. Any medium is made up of symbol systems which in turn are made up of coding elements.

multimedia courseware: Term defining high-level computerized learning programs which drawing on hypermedia methods give users interactive control over text, sounds, graphics, animation, still/moving pictures and computer software. Multimedia courseware goes beyond the so-called CALL applications in so that latter does not necessarily presuppose the use of external or pasted video resources as a source of learning input.

OCD system: Optical Disc Corporation's 610-A recorder used to transfer video material from videotape to videodisc.

oral sample: Performance measure in which the subject is requested to give an up to three-minute speech (description or narration) based on the experimental video segments.

overall oral performance score: Composite score obtained by applying four rating scales (Effectiveness, Accuracy, Organization, and Fluency) to the transcript of the oral sample recorded by the subject.
"Practicing Spoken French" (PSF): Multimedia courseware package created for this study.

Semi-direct tests of speaking ability: Tests which elicit active speech by means of tape recordings, printed tests booklets, or other "non-human" elicitation procedures. They differ from direct speaking tests in that the latter includes procedures in which the examined is asked to engage in a face-to-face communicative exchange with one or more interlocutors (Clark, 1979).

simultaneous measures: Measures stored by the computer during the subject's practice with each one of the instructional units. The measures include total and detailed times on task, oral sample length, "A game"/"A quiz" scores, frequency and length of subjects' branching with the "Très Utile" stack, and number of recordings.

SLA: Second Language Acquisition.

subtitles: Lines of dialogue or of descriptive material shown on a motion picture or TV screen by superimposing them on the scene.

subtitling: Independent variable that controls which type of video segment can be accessed by the subject according to these conditions:

-- *subtitles*: The subject has access to a subtitled video segment throughout his/her practice speaking lesson.
-- *no-subtitles*: The subject has access to an identical segment but without subtitles.

task level: Independent variable that decides the complexity -- *higher or lower* -- of task to be performed by the subject. The variable is a function of the amount and depth of information.
from the video segments that subjects have to process in order to accomplish the required task.

**task type:**
Variety of instructional unit, description or narration, which subjects can select for their speaking practice.

**transactional tasks:**
Tasks used to elicit spoken accounts of static (e.g., description), dynamic (e.g., narration), or abstract (e.g., opinion-expressing) relationships.

**video segment:**
Each of the four video excerpts, which pressed onto a videodisc and accessed from the computer screen, constitutes the source of the subject's oral output.

---

**Assumptions and Limitations**

The following assumptions were made in order to conduct this study:

1. Second-language oral communicative practice has as its primary goal to provide training in processing and conveying aural and/or visually-based information effectively.

2. Spoken accounts of static (description) and dynamic (narration) relationships are valid measures of oral communicative performance.

3. Subjects will participate cooperatively and will produce oral samples that best reflect what they have learned from their practice with multimedia courseware.
4. Practice with multimedia courseware may provide subjects with opportunities\(^7\) for producing meaningful output.

The following limitations must be taken into account before any attempt to generalize from the results of the study:

1. Because of the short duration of the experiment, any effects of the two treatment variables can only be generalized as the effects of short-term practice under the treatment.

2. The semi-direct testing procedures utilized in the experiment cannot measure general speaking proficiency. The somewhat artificial kinds of language tasks performed by the experimental subjects cannot be taken as direct indicators of speaking ability in real-life situations.

3. Any study of oral communicative practice with multimedia courseware should take into account the key role played by the video segments and the tasks chosen to sustain that practice. Because the study evaluated performance on only two tasks and four video segments, results should be generalized cautiously. Replication of the study with different tasks and segments is advisable.
Overview of the Dissertation

A description of the study briefly introduced in the preceding pages will unfold in five subsequent chapters as follows. Chapter 2 will begin by reviewing theory and research relevant to the independent variables (subtitling and task level), and the dependent variable (oral communicative performance). The chapter will next review literature on the assessment of oral communicative ability in second/foreign language, and it will conclude with the review of research on five CALL issues indirectly addressed by the study. Chapter 3 will explain the procedures used in both the pilot and the actual studies, discuss the validation of the scoring instruments, and present the means by which data were collected and computed. Chapter 4 will describe the design and content of the courseware package "Practicing Spoken French" created by the researcher for use in the experiment. Chapter 5 will present the results of the experiment while Chapter 6 will discuss the implications of those results and provide recommendations for further research.
CHAPTER 2

REVIEW OF THE LITERATURE

Introduction

This chapter reviews literature relevant to the independent variables (subtitling, task level) and the dependent variable (oral communicative performance) identified in this study. The chapter also reviews research dealing with five CALL effectiveness-related issues about which this study may provide useful insights.

CALL research on the independent variables of the study is far from impressive. Few studies have addressed either the subtitling or the task level variables, and to my knowledge, no study has considered the two variables conjointly. There are two probable reasons for this lack of attention: 1) the manipulation of such variables in the way proposed in this study requires a sophisticated software able to provide meaningful speaking
practice; 2) the production of such software is bound to be costly and time-consuming.

Regarding the subtitling variable, most of the foreign/second language studies on the issue have concentrated on the effects of subtitles on comprehension, reading, or listening. None has investigated the effects of subtitles on productive skills, such as speaking. These studies are "impressionistic" in nature and frequently characterized by their lack of systematic empirical procedures.

As concerns the task level variable, the few CALL studies which have addressed the effects of levels of processing have focused on prose reading practice (Pederson, 1985), rather than on speaking communicative practice. Most of the research on task level and task structure draws from fields, such as instructional media or scientific domains, which have a strong base in cognitive psychology. However, this avenue of research has begun to attract the attention of CALL specialists, as evidenced by recent discussions about the types of CALL activities which might be appropriate in generating learner's "talk" with computer programs. Esling (1991), for instance, suggests the adoption of Brown and Yule's (1983a) task-based, discourse-oriented framework to generate/evaluate CALL speaking production. According to Esling, for each of Brown and Yule's task types\(^1\), CALL equivalents could be identified and divided into two phases: 1) a
phase for receptive practice and manipulation of the perceptual aspects of language (input), and 2) a phase for practicing oral and written production during the CALL activity (output).

With respect to the dependent variable, oral performance, a large number of studies have addressed the assessment of communicative abilities in second/foreign language. This testifies of the "effervescence" of a field struggling to define the communicative and creative components of proficiency, and to establish the type of tasks and parameters by which a language test can be defined as "communicative" or "authentic."

Referring to research on CALL effectiveness, the studies carried out during the last decade reflect an incipient, but promising trend moving away from comparing CALL vs. teacher/audiotape/TV, and focusing on the specific conditions under which use of computer programs may result in second/foreign language learning.

This review of the literature begins with a discussion of the three theories which may provide a theoretical framework for the main independent variable, subtitling. This is followed by a description of existing studies on the effects of this variable. Right after, the chapter examines how task level and task structure have been shown to be significant in research in both instructional media and the sciences. Next, certain tests of communicative performance in foreign/second language are reviewed from
the viewpoint of their elicitation tasks and performance criteria. Based on that review, some conclusions are offered about what should be observed when developing, or selecting, instruments to assess direct, or CALL-mediated, oral production. Finally, the chapter offers a condensed review of research on the five CALL issues which are indirectly addressed in this study.

**Subtitling**

**Theoretical Basis**

Three types of theories, generated and/or applied by research on cognitive psychology and mass communication, could be invoked to explain the potentially superior performance of students exposed to fully duplicating intralingual subtitles\(^{(2)}\): serial--parallel processing, bottom-up--top-down processing, and between channel redundancy.

**Serial--Parallel Processing**

As psychologists emphasize (Anderson, 1980), the mind’s overall processing capacity is enormous, but the limit of its conscious attention at any one moment is comparatively small. Mental tasks that have been overlearned to the point of becoming automatic require very little space. Those tasks that must receive conscious guidance and control occupy a much larger area of brain processing space. The reason for the latter is that the brain must attend to each aspect of the process or action one step at a
time in a "serial" order. In the case of routine behaviors, attention to individual components is unnecessary, and the brain has the additional capacity needed to process several activities simultaneously in "parallel" fashion.

The use of serial and parallel processing is not mutually exclusive, both are used as practice and familiarity make more efficient use of brain space. Therefore, it could be expected that, with the appropriate amount of practice, learners could switch from a serial to a parallel, and more effective, processing of aurally (narrative) and visually (subtitles) cued information.

**Bottom-Up--Top-Down Processing**

As psychologists also suggest (Madden & Nebes, 1980), the brain can approach a subject from two different directions, bottom-up or top-down.

In a bottom-up approach, the brain begins with externally received stimuli, and analyzes that information to arrive at its final interpretation. In a top-down approach, the brain initiates processing with some general knowledge of the subject and proceeds to relate information it already has to the new data it receives.

The bottom-up and top-down distinction is commonly used in reading research (Pederson, 1985) to describe how readers interact with text. Bottom-up processing can begin with letter shape recognition or morpheme
recognition, and is usually associated with analysis and decoding. It is particularly evident in unskilled readers who either lack adequate background knowledge to rely on context for meaning or lack adequate knowledge of graphic language to proceed at a rate that encourages meaningful encoding. In order for top-down processing to take place, readers must have adequate schemata available to match the story grammar. If this real knowledge is not available, or if the text present novel information, they may resort to bottom-up processing as a back-up strategy to compensate for the inability to process meaningfully.

The top-down-bottom-up construct could also be applied to listening comprehension where the bottom-up processes would be associated with phoneme and word recognition and the top-down processes with semantic encoding of the aural message.

Performing a short cutting function, a coding element such as subtitling could facilitate the bottom-up processing of the aural message thus permitting the listener to focus on the top-down processing of that message. In order to promote top-down processing, coalescence of subtitling and perception of task difficulty would be required because, as Salomon (1979) notes:

... the perception of the task to be performed determines the kind of information one wishes to extract from a coding message, thus entailing a "top-down" type of
processing. In the service of extracting that information, codes are encountered and mental skills are brought to bear on them to transform them from external to internal representations, where meaning is assigned to them. This is a "bottom-up" process. (p. 112).

**Between-Channel Redundancy**

The third type of theory invoked to explain the potential outperformance of students exposed to subtitled video is grounded in research on the effects on channel redundancy, between-channel or single-channel, in audiovisual presentations.

The reason advanced by between-channel redundancy advocates for expecting superior performance from audio-print materials is that redundant information presented in two sensory modalities, aural and visual, can be conceived of as two separate learning trials. Thus, since the viewer receives two instances of the same material, more learning is predicted. In this respect, Singer (1980) observes that:

> Although television clearly has an advantage when issues of recognition-memory are at a premium, situations in which one must voluntarily call up material that has been previously learned are maximized if one has in effect stored this material in two locations - in the verbal encoding system and in the visual system. (p. 59).

Justifying his observation, Singer adds that pictorial information stored in the right brain hemisphere is hard to retrieve if not actively
processed, and he suggests that verbal labels could facilitate optimal retrieval.

Following subtitling proponents, between-channel redundancy could be explained by theories such as energy-summation, multiple-trace, or cue-summation. The energy-summation theory predicts intersensory facilitation when audio and visual information occur jointly (Nickerson, 1970). The multiple-trace hypothesis predicts superiority of transmitting information in more than one modality due to multiple strength of memory traces (Wickelgren, 1970). The cue-summation theory states that cues are provided from one channel to the other when audio-visual redundancy occurs (Adams & Chambers, 1962).

Referring to the last of these theories, Hartman (1961) contends that the greater the number of cues, the more opportunities there are to make discriminations between one stimulus complex and another. He concludes that the summation of clues is likely to facilitate learning when: (a) redundant information is transmitted simultaneously through print and audio channels; and (b) when cognitive association facilitates linking information units, as in a successful verbal label for an ambiguous picture. Interference is likely to occur when: (a) information between channels is unrelated or when the difficulty or presentation rate is such that alternation of attention
between channels is not possible, and (b) when contradictory cognitive relationships are present between information units in either channels.

Other paradigms, such as Broadbent's (1958) single-channel perceptual system, oppose the between-channel redundancy model and put forward arguments against the use of subtitles (Mondfrans & Travers, 1964). According to Broadbent, informational inputs from only one sensory modality at a time are said to have access to the higher brain centers. Thus, information in one channel would interfere with that in another. Broadbent's theory is somewhat debatable, because, as Ortony (1978) and Grimes (1990) contend, channel interference is likely to occur only when the visual channel activates memory traces contradictory to those activated by the verbal channel.

Empirical Findings

First-Language Research on Subtitles

Instructional specialists in fields as different as hearing-impaired education or broadcasting have seen in the use of subtitles an effective way of improving viewer information processing capacity.

A number of experiments on the effectiveness of subtitles have been concerned with the teaching of reading skills to hearing-impaired children. Baker, Downton, and Newell (1980), for instance, report on a experiment in which partially and fully duplicating methods of subtitling a British soap
opera were used. Results, which confirmed Levin's (1979) theory about the eye-vocal-voiue-span, indicated that both prelingually deaf subjects (with heavy hearing deficit) and postlingually deaf subjects (with mild hearing deficit) preferred fully duplicating subtitles (almost verbatim representation of the shooting script).

The issue of subtitles modification has also been addressed in a study by Hertgoz, Stinson and Keiffer (1989). In this study, two technical versions of a film, captioned at 8th- and 11th-grade readings levels, were shown to 32 hearing-impaired college students. Data from a comprehension test were analyzed with a four-factor experimental design to determine effects of instruction, level of captioning, test type, and subject reading ability. Results indicated that while high- and low-level reading students benefitted from instruction when they viewed 8th-grade level, partially duplicating captions, only the high-level reading students benefitted from instruction when they viewed the 11th-grade level, fully duplicating captions.

As concerns broadcasting research, a number of experiments have tested the relative efficiency of the various communication channels as information transmitters. Generally speaking, experiments comparing the combination of the audio and print channels to either the audio or the print (Elliot, 1936; Hartman, 1961; Hsia 1968; Grimes 1990), have found
advantage for the audio-print combination. While presenting evidence that between-channel redundancy can reduce error and facilitate information processing, Hsia (1961) observes that one channel may provide cues for the other, provided that the processing capacity limit is not exceeded.

However, other researchers, such as Reese (1982), have also found that adding redundant print information does not increase efficiency in the processing of broadcast news information. Reese’s finding seems to support Broadbent’s theory that when two channels are included in a television news story (print and audio), viewers must split attention between the two with a resulting loss of information. But, Reese’s finding would maybe have been different if, as the researcher observes, “The pace and complexity of visual material could be varied and more time provided to digest printed information.” (p. 81). Or, in other words, if viewers have had control over both channels.

Second-/Foreign-Language Research on Subtitles

The above theories and findings about subtitles also seem to apply when subtitled materials are intended for the second/foreign language learner, as evidenced by research conducted in three sub-areas: text-sound association, reverse subtitling, and subtitled video.

Text-Sound Association. In the mid-sixties, research within the audio-lingual tradition addressed, with varying success, the issue of how the
written word was an aid for training in discrimination of foreign sounds and for evoking oral responses at the beginning levels (Muller, 1965; Estarellas & Reagan, 1966).

In Muller's (1965) study, for instance, 144 tenth-, eleventh-, and twelfth-grade American students of Portuguese were randomly assigned to an experimental and control group. Subjects in both groups were tested on their pronunciation in Portuguese, following a period of instruction using intensive audio-lingual techniques. The pronunciation test consisted of twenty-one sentences in Portuguese, which were suggested by the picture-cue sheets with which each student was provided. All students received exactly the same instruction for the same amount of time. The only variable was the printed Portuguese words on the picture-cue sheets used by the experimental group. Results of the pronunciation tests indicated that the control group consistently scored higher than the experimental group. Muller "categorically" concluded that "exposure to written words in Portuguese during the early stages of learning that language will have an adverse effect upon pronunciation proficiency in Portuguese of English-speaking students." (p. 413).

In Estarellas and Reagan's (1966) study, ten students were randomly selected from a Level I Spanish class and subjected to intensive training in discrimination and pronunciation of Spanish sounds and their written
equivalents. The material was a programmed self-instructional text coordinated with tape recordings and included levels on vowels, consonants, linking, stress, intonation, and syllabification. A control group began the Audio-Lingual Materials (A-LM) in the conventional manner (no written material). At the end of the self-instruction period, the students of the experimental group moved into the A-LM, and they quickly overtook the control group in dialogue mastery and pronunciation, as proved by achievement tests administered to both groups. In light of the results, and contrary to Muller (1965), Estarellas and Regan, concluded that "when presented on a self instructional basis, the students learn the phonemic-graphic associations faster and achieve a higher degree of mastery in learning the dialogue sentences in a beginning FL course." (p. 181).

Abandoning the research perspective of the audio-linguists, Berwald (1979) showed how teacher-made captions to accompany slide and film presentations in the foreign language can be extremely useful in helping students to understand an accompanying taped narrative or film soundtrack. According to him, the subtitled materials he created for his French conversation class seemed "to help students develop their language skills." (p. 376). Interestingly, Berwald was one of the first to mention the possibility of adding and removing subtitles -- an option recently made possible by broadcast videotext technology.
Reverse Subtitling. A number of studies have investigated the use of reverse subtitling and its effect on listening comprehension. In a study involving 370 upper-level elementary school Canadian students of French, Lambert, Boehler and Sidoti (1981) explored novel uses of scripts and dialogues in media presentations: A TV screen was used as an adjunct to radio in such a way that the dialogue presented aurally (either in first language or second language (thereafter, L1 or L2)) was matched word-for-word with a script presented on the screen (either in L1 or L2). Three main combinations emerged from the various input conditions: (a) "Standard Subtitling" (dialogue in L2 and accompanying script in L1), (b) "Bi-Modal Input-L2" (dialogue and script in L2), and (c) "Reverse Subtitling-L2" (dialogue in L2 and script in L1). Subjects under the different conditions were tested for comprehension, spelling, contextual meaning, and phrase form of the spoken dialogues. Results indicated that the Bi-Modal Input-L2 condition was, even at the novice level, particularly beneficial for remembering phrase form in L2. The same condition also proved to be globally more effective than Reverse Subtitling-L2 for advanced L2 students. According to the researchers, the Bi-Modal Input-L2 modality, could have greater promise as a learning/studying device because it provides subjects with the opportunity of mapping L2 dialogue directly onto L2 script, with the interplay of script and dialogue possibly working both
In a replication of the above study, Holobow, Lambert and Sayegh (1984) tested three ways of combining visual and auditory inputs of messages: "Reverse Subtitling," "Bi-Modal L2 input," and "L2 script-only." As in the previous study, the experimental subjects were Canadian elementary school pupils with advanced training in French (L2). Matched sub-groups were assigned to particular treatment conditions that ran for a ten-week period. As predicted, the Bi-Modal L2 input condition appeared to be particularly suited for enhancing comprehension skills of advanced L2 learners. Agreeing with the advocates of the between-channel redundancy model, Holobow et al. concluded that the double modality input appears to facilitate the processing of information better than the single channel L2 (either auditory or visual).

**Subtitled Video.** Arguments in favor of the use of subtitled video in second/foreign language learning also abound. Jung (1990), for instance, drawing on Bullard's (1985) theory about word perception, states that "Subtitles might serve as advance organizers..., which diminish the decoding load placed upon the learner by the unrefined audio signal of authentic materials." (208-209); and he adds that "If native speakers are in need of visual support to understand spoken announcements, should it be
denied to foreign-language learners?" (p. 209). In the same way, Ellsworth (1992) contends that subtitled video may "help people who are insecure in their second language to use it, interact with it, and expand it." (p. 24). Ellsworth provides interesting suggestions based on his use of subtitled American video for Spanish audiences. For instance, he explains how to use subtitled video to teach grammar, vocabulary, slang/idioms, cultural expectations, and social etiquette. He also explains how he exploits, through the use of a decoder, different types of presentations: video with subtitles and audio, video with subtitles, video and audio, audio only.

In spite of these arguments and pedagogical suggestions, only a few studies have tested the use of subtitled video materials to teach English as a second language (Jensema & Fitzgerald, 1983; Price, 1983; Vanderplank, 1988; "Closed-captioned TV," 1991), or even the so-called "less commonly taught" foreign languages (Garza, 1991).

Price's (1983) 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. Some 500 students of over 20 native language backgrounds participated. Half of the subjects saw four video excerpts with subtitles; the others saw the same video segments without subtitles. Half of each group was exposed to two
viewings; the other half to a single viewing. Measures concerning viewing comprehension, as well as extensive background and sociolinguistic data, were obtained for each student. The effects of these variables along with segment preference, repeated viewing, and order of presentation were evaluated. Results indicated that "viewers, regardless of educational level or language background, benefitted significantly from captioning, even with only one viewing." (p. 8). Price's findings seem to run counter to the opinion of those for whom subtitling is a sociologically bound phenomenon. Providing to such opinion, Vöge (1977) states that "sub-titling is more often preferred as one is younger or more highly educated." (122).

In Vanderplank's (1988) study, fifteen European and eight Arabic learners of English, between high-intermediate and post-proficiency level, watched nine hour-long sessions of BBC general output television programs with English subtitles. Subjects were asked to provide detailed feedback on the following aspects: language gained from the programs, reactions to the subtitles, strategies used in exploiting the subtitles, levels of anxiety, comprehensibility of the sound and text, and programs themselves. The subjects were also asked to accomplish a number of language-oriented post-activities connected with the programs (e.g., acting out dialogues and scenes from situations seen on the programs, writing lists and explaining differences, and group discussions). Results of the reports basically
indicated that subjects found the subtitles beneficial to their language development. 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. Vanderplank also suggested that subtitles may be of limited value at beginner and low-intermediate levels since, at those levels, the language of the programs may be, even with textual support, beyond the learners in terms of grammar, lexis, and speed of delivery. However, according to him, subtitles can be beneficial for post-intermediate-level learners by "turning" television programs into sources of comprehensible language input.

In the 12-week study conducted by Neuman ("Closed-captioned TV", 1991), 129 seventh- and eighth-grade students from Asian and Hispanic backgrounds were randomly assigned to one of four groups: (a) viewing closed-captioned television, (b) viewing television only, (c) reading and listening to the text, and (d) 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 90 target words. Students in the closed-captioned group scored higher than those in any other group.
Concerned with the generalizability of findings of L2 research on subtitling to the "less commonly taught" foreign languages, Garza (1991) conducted a comparative study using Russian and English (ESL) as target languages. Subjects in the study were 40 third- and fourth-year students of Russian at two American universities, and 70 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, respectively lasting 16 and 18 minutes. Subjects in the Russian and American classes were randomly divided into two groups so that half of them could see the test segments (with subtitles) and half the control segments (without subtitles). Two corpora of data were collected: 1) scores of all students on ten content-based comprehension checks in both languages with and without captions; and 2) coded transcripts of the retelling, by five randomly selected subjects, of the content of a chosen video segment.

Results of the first corpus of data indicated an overall gain in correct responses for subtitled versions of the ESL and, particularly, the Russian video segments. Results of the second corpus of data also attested to the positive impact of subtitles on the subjects' language facility. Subjects consistently demonstrated a greater propensity to use the original lexicon of the video segments when viewed with subtitles.
Virtually, all of the above studies have shown the usefulness of subtitles for bridging the gap between the L2 learners' competence in reading and listening. None of them, however, have addressed the 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 a proper manner." (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.

**Learning Task**

According to Salomon (1979), the essential research question in choosing the combination of coding element and task variables to investigate is whether or not the mental activities that the coding element activates are task-relevant. In the same way, Pederson (1987), based on the results of her task-oriented reading study, observes that "students tend to assume that certain ways in which CALL is presented deserve special attention, depending upon their perception of the task they are expected to perform." (113).
Two ways of improving the relevancy of the learning task may be to increase its complexity (task level), and to raise its formal elaborateness (task structure). Support for this concept is provided by cognitive theory and research in instructional media and in the sciences.

**Task Level**

**The AIME Model**

Salomon suggests that learners' preconceptions about the difficulty of the medium (e.g. "easy" television vs. "difficult" computer) may be influenced by learners' perceptions of the learning task.

According to Salomon's (1981, 1984) Amount of Invested Mental Effort (AIME) model, the amount of mental effort invested in processing information provided by a medium depends on two types of perceptions. The first type, Perceived Demand Characteristics (PDC), is the degree to which a medium is perceived to pose demands. The second type, Perceived Self-Efficacy (PSE), deals with the amount of mental effort put forth. One's perceived self-efficacy toward a medium is high if one perceives oneself to be quite capable of obtaining information from that medium.

In a vein similar to Salomon's, cognitive psychologists (Anderson, 1980; McDonough, 1981; Horton & Mills, 1984) contend that the quantity
and difficulty level of a learning task is an important factor contributing to the retention and recall of information presented or required by that task.

**Application of the AIME Model**

Research on instructional media has investigated the AIME model in relation to three different variables: the learners' past experiences (Gagné & Glaser, 1987), the learning task (Krendl & Watkins, 1983; Salomon & Leigh, 1984), and the characteristics of the lesson (Berlyne, 1960).

Focusing only on the AIME/learning task studies, Krendl and Watkins (1983) provide evidence that increasing the amount of effort required in a task improves learners' performance. Sixty students from a public school in Michigan were assigned to one of four treatment groups representing the cells of a 2 x 2 factorial design. Half of the participants were informed that the program was to be "educational," while the other half were told that it was simply intended "to entertain." Similarly, half of the participants were given the opportunity to stop the presentation at any time, controlling the pace of the videotape, while the others were not given that option. Treatment groups were based on the manipulation of viewing set (either educational or entertaining), and viewing mode (either stopping or no-stopping). The visual input was provided by a 15-minute episode from the educational series "Thinkabout." This series, designed to teach basic problem-skills, presents live actors in fairly realistic social situations. The
series was chosen because it could be considered either educational or entertaining, depending upon one's frame of reference. The educational context was designed to impose high demands on the viewer and to enhance more active viewing, more elaborations and inferences, while the entertainment context was intended to facilitate viewing in a very low-demand context. A set of measures were devised to assess processing beyond the level of simple recall. The measures included the recall of additional detailed program material, and the interpretation or subjective commentary on events in the program. Results offered significant evidence for differential processing between viewing groups. Subjects who were assigned to the educational viewing set (high-level) gave more analytical statements and abstract reconstructions than subjects on the entertaining set (low level). In light of these results, Krendl and Watkins concluded that "the process of attending to and learning from television ... becomes a function not only of the messages sent, but of the perceptual set by which the messages are received and interpreted." (p. 211).

Salomon and Leigh (1984) also tested the AIME model on a series of three studies with Israeli and American children. In the second of the studies, eighty-seven sixth graders from a Jerusalem school were assigned to the cells of a 2 (TV vs. Print) x 2 (manipulated PDC) factorial design. Half of the children were shown a TV story, "The Violin," and the other half were
given the comparable version in print to read. Within each condition, half of the children were instructed to watch or read the story "for fun" (low perceived demand characteristics: PDC-FUN), while the other half were instructed to watch or read the story "to see how much you can learn from it" (high perceived demand characteristics: PDC-LEARN). Post-test scores consisted of a questionnaire measuring reported AIME, and tests of recall (five items) and inference making (five items). Results of the reported AIME questionnaire indicated that more mental effort was expended in both the TV and the print stories under the PDC-LEARN condition. While manipulation of PDC did not have significant effect on the recall measure, it did had an effect on the inference measure. The PDC-LEARN condition lead to significantly increased inference-making scores for the TV viewers who provided superior responses when asked about motives of the story's main characters and about implicit causes of the story events. In light of these findings, Salomon and Leigh concluded that "when the viewing task is perceived as more demanding, more effort becomes invested, abilities are mobilized, and inference generation is improved." (p. 132).

**Task Structure**

If theory and research seem to agree about how perception of difficulty of task affects performance, there is also theoretical and empirical
evidence that performance is equally affected by the formal organization of the task.

According to Bruner (1977), the transfer of learning, which is at the heart of the educational process, is dependent upon the structure of the subject matter, or the way content elements are interrelated. He observes that "knowledge one has acquired without sufficient structure to tie it together is knowledge that is likely to be forgotten." (1977, p. 31).

In the same way, Anderson (1980) contends that retention and recall of information is also a function, among others, of the degree to which the structure and the organization of the material is compatible with the individual's cognitive network.

Similarly, Reigeluth's (1983) Elaboration Theory maintains that the development of stable cognitive structure is best achieved by presenting content in a framework of a single, top-down organizing structure.

Underscoring psychologists' viewpoints, researchers on instructional design (Shavelson, 1972; Reif & Heller 1982; Eylon & Reif, 1984) have suggested and/or proved the correspondence between the organization of instructional materials (content structure) and the resulting organization of the internalized knowledge (cognitive structure).

In their 1982 article, Reif and Heller observed that "students are given little help to integrate their accumulating knowledge into a coherent
Eylon and Reif concluded that higher levels of hierarchical organization of knowledge "can appreciably facilitate educationally important recall and problem-solving tasks ..." (p. 39). Eylon and Reif's study illustrates the utility of an approach that strives to formulate information-processing models for enhancing learning performance, and which deliberately induces learners to act in accordance with those models.

Studies such as the above seem to indicate that the modelled structure of the content should be taking into account in organizing and sequencing computer lessons. Therefore, it would be interesting to investigate how the structure of tasks in multimedia programs for foreign languages affects performance. This investigation would be fully pertinent because, as Stern (1983) points out, "the experience, research, theory and discussion on educational technology outside the field of language pedagogy are prerequisites to an understanding of the application of technology to language teaching." (p. 445). Also, the investigation would be particularly appropriate at a moment when opponents of multimedia technology are arguing that being "lost in hyperspace" --being disoriented by associative complexity-- is a fundamental obstacle to learning by exploration using hypermedia tools (Nielsen, 1990).
The Assessment of Oral Communicative Competence in Second/Foreign Language Learning

The assessment of learners' ability to speak in a foreign language is an unquestionable challenge to both classroom teacher and test developer. During the height of audiolingualism, various testing methods were designed to evaluate L2 speaking skills. These tests concentrated on the evaluation of such discrete linguistic items as stress, intonation patterns, pronunciation of isolated sounds, or correct oral production of isolated grammatical features of the foreign language. As Stern (1983) would say, developers of such tests were unaware that they needed input not only from psychometrics but also from linguistics.

However, as the communicative language teaching movement has gained momentum during the last two decades (Savignon, 1972; Widdowson, 1978; Brumfit & Johnson, 1979; Robinson 1981; Rivers, 1986, 1887), language testing has espoused the views of that movement (Morrow, 1979; Porter, 1983), and a wealth of publications dealing with topics such as test reliability and validity (Ollier, 1979; Messick 1980; Palmer, Groot & Trosper, 1981; Alderson, 1988) has developed.

Within this new perspective, design of language tests invariably takes into account the following elements: the "communicative" theoretical framework that guides the development of the stimulus materials, the tasks
to which the test-taker must respond, the performance criteria, and the scoring procedures. Among those elements, tasks and criteria are paid particular attention because their selection determines the degree to which language generated by tests reflects, at least partially, "real life" language, and the extent to which tests' scores are a reliable measure of language performance.

Oral Communicative Tests: Tasks and Performance Criteria

An overview of some relevant tests of oral communicative competence reveals the variety of tasks used by those tests to elicit the appropriate speech samples from the test-taker, as well as the variety of criteria by which those tests evaluate test-taker's performance. This variety, in turn, reflects the "multiplicity" of the communicative competence concept and the schools of thought which have contributed to it.

Savignon's (1972) test of communicative competence for beginning college students of French includes four tasks (discussion, interview, reporting, and description) and ten performance criteria varying from "effort to communicate," "amount of communication," "naturalness and poise," to "comprehension by the native speaker," "comprehensibility," "fluency," and "accuracy."

Savignon's test takes into account the psychological implications of the artificiality of the language classroom and the nature of the practice
afforded to the language learner. For this reason, as, Canale and Swain (1980) observe, Savignon's performance tasks, have more face validity with respect to communication skills in that such tasks correspond more directly to normal language use where an integration of these skills is required with little time to reflect on and monitor language input and output. (p. 34)

Unlike Savignon's instrument, the tests designed by Clark (1972), Mothe (1975), and Buscaglia and Polly (1981) consist of only one task (interview or information-getting questions) measured by performance criteria which focus on the linguistic rather than on the communicative abilities of the testee (e.g.: fluency, pronunciation/intonation, grammar, vocabulary, comprehension, comprehensibility).

Bartz's (1979) test of communicative competence for intermediate-advanced levels, includes three optional tasks (description, interview, and simulation) and four scoring criteria (fluency, comprehensibility, amount of communication, quality of communication, and effort to communicate). Contrary to Clark (1972), who recommend that communicative and linguistic criteria not be mixed in the evaluation of communicative ability, Bartz contends that it is difficult to totally separate the two criteria, "as the linguistic quality of an utterance can influence comprehensibility, the basic communicative criterion." (1979, p. 13). According to Bartz, the question
that must be dealt with in the evaluation of the speaking skill is to what degree discrete errors interfere with the intended message of the speaker.

In their investigation of the learning strategies used by learners of English as a second language, O'Malley et al. (1985) utilize an oral performance instrument similar to Bartz's. The instrument consists of an oral report task and four performance criteria including delivery (volume and pace), appropriateness (choice of words and phrases for a class presentation), accuracy (phonological, syntactic, and semantic), and organization (coherence and cohesion).

Only a few tests have taken primarily into account the sociolinguistic or strategic components of Canale's (1983) communicative competence model. Carroll's (1980) oral evaluation procedure, for instance, consists of a talk-discussion task, and four performance criteria including grammar, amount of information, complexity, and sociolinguistic adequacy. Similarly, Henrie and Monfils' (1985) test utilizes a role-playing task assessed against performance criteria such as grammar, comprehensibility, sociolinguistic adequacy, and strategies.

Portal's (1986) French test, developed for use in a survey of foreign language performance of 13-year-olds in the United Kingdom, consists of four speaking tasks (structured conversation, role playing, description, and report of pictorial-based events) and three performance criteria (amount of
information, appropriateness, accuracy) measured by three five-point scales. Portal's test exemplifies of what Bachman (1990) defines as the interactional/ability (IA) approach to "authenticity," in that it takes into account the interaction between the test taker, the test task, and the testing context. The tasks in Portal's test reflect the type of tasks subjects are likely to meet while in the foreign country. They are aimed at maximizing the student's opportunity to speak, within a 10-minute testing time, and producing sufficient French to allow a reliable assessment to be made.

A test falling into the category of semi-direct integrative tests is the Recorded Oral Proficiency Examination (ROPE) for various foreign languages developed by Lowe and Clifford (1980). The aim of the ROPE was to approximate as closely as possible the linguistic content and manner of operation of a live interview. The test consists of a series of tape-recorded questions in the target language to which the examinee replies orally, with responses tape-recorded for later evaluation. The test is scored and interpreted on the basis of the ACTFL/U.S. government Interagency Language Roundtable (IRL) general proficiency guidelines. A validation study conducted by the test developers showed a high correlation ($r = .90$) between examined ratings on the ROPE and on a concurrently face-to-face interview.
Another test belonging to the semi-direct integrative category is Clark's (1986) test of Chinese speaking proficiency. Developed after the ROPE, the test consists of four basic tasks including simulated conversation, picture-based discourse (giving directions, narration), response to printed questions (stating opinions), and role-playing. Rating criteria include appropriateness, register, style of speech, and various phonological features.

Either direct, semi-direct, or semi-direct integrative, the above tests utilize rating scale evaluations which according to some theorists (Corrigan & Upshur, 1982) may lead to a method/trait confusion in the testing of communicative competence as opposed to linguistic competence. Therefore, alternative tests have been developed which use discrete point evaluation procedures such as the Active Communicative Competence Test (ACCT) created by Politzer and McGroarty (1983). The ACCT involves three tasks: 1) giving map directions in such a way that the listener is guided to the desired destination; 2) inviting somebody to a party on the basis of relevant information provided in pictorial form; and 3) reporting an accident described by a series of cartoon pictures. The reliability of the ACCT, calculated according to the consistency measure Cronbach's $\alpha$, is quite high (.93), and so is the correlation of the test with other linguistic competence measures such as the Test of English as a Foreign Language.
(TOEFL) (.47), and the Comprehensive English Language Test (.58). In the view of the developers, the ACCT provides an illustration of how traits of communicative competence can be measured by the same discrete point method as linguistic competence.

**Some Working Conclusions**

The overview of tasks and criteria used by the above tests leads to a number of conclusions which should be taken into account by researchers and/or developers confronted with the selection and/or creation of instruments to assess L2 oral production in general, and CALL speaking practice in particular.

1. Task-oriented activities may serve as integrative tests of learner’s competence to produce and understand actual communication in the oral mode. Expanding the variety of task-types used in oral tests can provide the test-taker the chance to structure long turns and produce a richer structure of "talk." Varying the tasks can also contribute to the objective quantification and assessment of the "real-language" which may result from such tasks.

2. The criteria to evaluate oral performance tasks should mirror what Spolsky (1989) names the "minimally relevant dimensions" (p. 147) of communicative competence rather than the components of a specific model (e.g.: Canale’s, 1983). Those criteria would provide a reliable indicator of the authenticity of the test tasks. As Bachman (1990) puts it,
If we could develop a means of classifying test tasks on the basis of dimensions, or factors that we abstract from authentic language use, we would be able to characterize the relative authenticity of a given test task in terms of its potential for generating an authentic interaction with the abilities of a given group of test takers. (p. 317)

3. Communicative proficiency should be defined in terms of component rather than unitary abilities, and thus, scores for these abilities should be separate rather than global.

4. Tests should not have to necessarily measure the entire range of competencies included in a given theoretical construct of communicative language ability. Selection of these competencies should depend on the aspects of communicative abilities one wants to measure. As Stevenson (1982) points out, "To state that there is only one proper goal of testing, or that a test with any other goal cannot be valid is to pursue rhetoric that has very little to do with scientific validity and validation." (p. 10).

5. Tests of communicative ability shouldn’t, in any case, overlook the linguistic component because if grammatical inaccuracies are not corrected, they may fossilize on an interlanguage that does not correspond to the language system used by native speakers of L2 (Canale & Swain, 1980). And also, because formal accuracy can help to approximate the mental representations that learners construct to the mental representations of native speakers (Rivers, 1991).
6. Neither direct nor semi-direct tests\(^{(6)}\) of speaking ability are superior to each other (Clark, 1979), and both types may have high scoring reliability, as found in the tests mentioned above (e.g.: Savignon (.95); Bartz (.99); O'Malley et al. (.85); Portal (.86)). Therefore, the choice of each one of them should be a function of its practicality.

7. Although discrete-point tests may have, as seen earlier, a high degree of face validity and scoring reliability, they should be disregarded when the materials required by the speaking tasks make the content analysis of the discrete elements of those materials (e.g. video) very difficult. In this case, criterion-referenced evaluations would be preferable.

8. The fundamental considerations in developing tests of language ability should be: 1) to match the appropriate testing methods to the various needs and purposes (Spolsky, 1986); and 2) to empirically assess tests' ability to measure specific language communicative competencies (Bachman, 1990).

Research on CALL Effectiveness: Five Issues

Effectiveness data base for CALL is limited and somewhat contradictory. Aware of this fact, CALL specialists (Dunkel, 1991) have urged researchers to engage in basic research efforts, and specify the issues that should be addressed by those efforts.
This review of research on CALL effectiveness will be limited to the issues on which the present study might provide some insight. The issues include: locus of control, time on task, learning environment, word processing, and attitudes.

**Locus of Control**

The belief that learners should be involved in choice-making when controlling learning has been widely held in many educational settings. Merrill (1975), for instance, indicates that instruction may increase in effectiveness by permitting learning control. Papert (1980), in his work with LOGO, argues that increasing control enhances feelings of self-efficacy and assists learners in taking independent responsibility for their own learning and behavior.

However, CAI research has shown mixed results for learner control. While results from some studies suggest that individuals learn more when given control over their own learning (Tennyson & Rothen, 1979; Gray, 1987), other studies demonstrate better performance when instructional decisions are made by the program (Gay, 1986; Robson, Steward, & Whitfield, 1988).

Gay's study (1986) investigated the effects of learner control and prior understanding in computer assisted video-instruction. 156 college students from an introductory biology course were randomly assigned to a program-
controlled or learner-controlled treatment. Under the program-controlled treatment, subjects had to complete the learning task in a specified order, and they controlled only the pace of the lessons. Under the learner-controlled treatment, subjects controlled not only the pace of the lessons but also the sequencing, presentation mode (video, audio, graphics, or text), amount of practice, and depth of study. Results indicated that subjects performed better in the program-controlled treatment than in the learner-controlled treatment. Results also demonstrated that learner control might be affected by the learner’s prior knowledge of the instructional task.

While studies such as Gay’s have shown learner control to be associated with poorer performance, the addition of some form of advisement to a learner-controlled lesson has been associated with greater achievement (Johansen & Tennyson, 1983; Shyu & Brown, 1992; Arnone & Grabowski, 1992).

Arnone and Grabowski’s (1992) study evaluated the effectiveness of variation in learner control on learning factual knowledge from computer-based interactive video instruction. The content was presented as a videodisc visit to an art museum. A posttest-only control design was employed with 101 school subjects randomly assigned to one of three experimental conditions (Program Control, Learner Control, and Learner Control with Advisement) or to a control group. Subjects in the Learner
Control with Advisement group were given the same control opportunities as the control group (sequencing, pacing, stopping, remediation and exiting). But, they were also given some "advisement" strategies that provided guidance and encouraged curiosity. For example, a student who decided to skip a section would receive this advisement: "Are you sure you want to end the lesson? This next section is very interesting. You might really enjoy it."

Posttest scores indicated that the Learner Control with Advisement treatment resulted in the greatest achievement.

CALL theory and research on the locus of control issue is also contradictory. While theorists have argued that one of the potential benefits of CALL is that it allows learners to "control" their learning experience (Underwood, 1984), studies have failed to prove evidence of the benefit of learner-controlled programs (Lahey, 1978). Recent studies, however, have shown that a midway position between learner- and program-controlled design may be beneficial. Robinson (1989), for instance, tested four variations of program-controlled feedback and student-controlled help. She found that the variation "combined program and student-controlled help" was more effective than any other three alternatives: 1) "no help, programmed correct answer"; 2) "total program-controlled help"; 3) "total student-controlled help." Robinson hypothesized that, under the learner control treatment, subjects did not diagnose their errors properly due to
their ineffective use of control options. On the other hand, under the total program control, subjects lose their motivation to work on the lesson.

Theoretical and empirical disagreements on the locus of control issue are even greater when it comes to instructional hypermedia. Against the cornerstone of instructional hypermedia, teaching by discovery and individual experience, there is mounting evidence that learners tend not to choose wisely when confronted with learner-controlled hypermedia programs (Jonassen & Grabinger, 1990).

**Time on Task**

In the past, CAI studies on the time factor (Orlansky, 1983) investigated the question of whether or not CAI programs can help learn more academic content in less time. Interest in this variable seems, however, to have diminished. Roblyer et al. (1988) suggest that decreasing learning time is not a practical concern particularly for school-age populations, since public schools have a set time for grade levels and courses. In the same way, Johnson (1985) stresses that the nature of the CALL tasks must be assessed for their appropriateness because, while the time saving factor is an important consideration, so too are the content and design of the CALL materials that generate language-using tasks.

Aside from savings concerns, the investigation of time on task may be useful in many respects. For instance, as Pederson (1987) observes,
simultaneous time on task measures taken by the computer may provide evidence about the learning strategies used by learners during CALL practice. Time on task may also provide information about how the locus of control affects the interaction of learners with computer-based instructional systems, interaction which has been proven to result in greater learning (Schaffer & Hannafin, 1986). Finally, in multimedia applications, measurement of time spent on the different stages of a program may provide interesting cues about how the processing of video stimuli, and related coding elements, affects performance on the subsequent video-based tasks of that program.

Learning Environment

In their discussion of the cognitive effects of CAI, Salomon and Gardner (1986), contend that the instructional medium itself does not affect learning and, thus, it shouldn't be taken as an independent variable. Subscribing to this contention, Johnson (1991), notes that "Research in [sic] the social aspects of computer use needs to move, then, from a search for the average social effects of the medium to a recognition of, and an accounting for, the complex social nature of computer learning environments." (p. 67).

A small number of CALL studies have addressed the "complex nature" of the learning environment by investigating two elements, the
learning site and the practice materials. With regard to the learning site, ethnographic observations from Lozano et al.'s (1985) study indicated that students using the computer laboratory had a more positive attitude toward the required practice than those who used the language laboratory.

As concerns the practice materials, ethnographers in Blomeyer's (1985), and in Lozano et al.'s (1985) studies noted that students had a very favorable attitude toward CALL lessons which they knew had been written for them. Such students appeared to have been highly motivated to learn using CALL "tailored" for them by their own teachers.

Word Processing

The literature in the field of word processing is replete with anecdotal evidence extolling its benefits. Hawisher (1987), for instance, notes that with word processing "writers can, insert, delete, and substitute text with an ease hitherto unknown" (p. 145). In the same way, Esling (1991) contends that "a reason for initiating learners to CALL with the use of word processing is the 'featural' clarity of the word-processing task." (p. 122).

The findings of a number of studies have lent support to the supposition that the quality of students' writing is affected in positive ways by the use of word processing. From their survey of eight studies examining the impact of word processing on the quality of Non English Speakers' (NES) writing, Roblyer, Castine, and King (1988) concluded that the effects
of word processing were positive. According to these researchers, students who perceive their writing to be better as a result of using word processing also appear to hold more positive attitudes toward the process of writing itself.

Examining the influence of word processing in a writing-as-process-oriented curriculum, Sommers (1985) compared the writing of a control group of 44 students who were receiving process-oriented instruction without using word processing to the writing of an experimental group of 35 students who were receiving the same type of instruction but were using word processing. Results indicated that word processing improved the writing skills of those in the experimental group. The writers in the control group increased their scores on a 12-point holistic scale by 0.46 points (SD = 2.88), whereas the writers in the experimental group showed an increase of 1.26 points (SD = 2.27).

Other studies, however, have failed to show a causal relationship between attitudes toward word processing and quality of written product. Neu and Scarcella (1991), for instance, investigated the attitudes of 54 undergraduate students toward writing using word processing in their ESL classes at the University of California at Irvine (UCI). Data from the 38-item questionnaire administered to the subjects indicated that they paid more attention to the grammar, vocabulary, and organization of their papers when
word processing than when writing by other means. According to the researchers, additional data would have been needed to determine "the interactive effect of students' better perceptions of their writing abilities, and the quality of their English compositions." (p. 182).

Further research on the issue should investigate the extent to which word processing may affect speaking skills when it is used to fulfill a prerequisite writing task on computerized oral practice programs.

**Attitudes Toward CALL**

A number of studies (Kulik, Bangert & Williams, 1983; Kulik & Kulik, 1986, 1987) have suggested that students hold positive attitudes toward using computers. In their meta-analysis of 101 computer-based education (CBE) classes, Kulik and Kulik (1986) found that of the 13 studies that examined student ratings of the quality of college-level instruction, 9 studies found more positive attitudes toward instruction in the CBE class.

There is a controversy, however, about whether or not attitudes toward computer are a function of the degree of "computer literacy." Based on their meta-analysis of 200 studies which analyzed the impact of microcomputer-based instruction on teaching and learning, Roblyer et al.'s (1988) observed that students inexperienced in using computers previously reacted more favorably toward computers than did students who were familiar with their operation. On the contrary, Ahmad, Corbett, Rogers, and
Sussex (1985) suggested that nonnative speakers' attitudes toward CALL work improves if students increase their computer skills.

Students attitudes toward the subject matter of computerized learning are not seemingly affected quite as positively as a result of using computers. In the Kulik and Kulik meta-analysis (1986), only 6 of the 15 studies examining the effects of computer-based instruction reported that student attitudes in CBE classes were more positive toward the subject matter than in the traditional classes. In the same way, in her computer-assisted reading study, Pederson (1985) concluded that "the subjects demonstrated a moderately positive attitude toward their computer reading practice experience." (p. 177). The one exception to this finding may be word processing which, as found in the studies mentioned earlier, seems to motivate students to write and revise more, and to share their writing with others.

Overall, if student attitudes toward the computer, and, to a lesser degree, toward subject matter, seem to be positive, insufficient data exist to indicate that better attitudes have any impact on achievement or on drop-out rate (Roblyer et al., 1988). Computers do not seem to have an impact on students' motivation to learn the subject matter even though students may report that they "like" to use computers. Maybe the key for that motivation
lies in the design of CALL programs that, going beyond the electronic book, may increase learners' perceived self-efficacy through learner-tailored tasks.

Summary

This chapter has reviewed theoretical and empirical literature relevant to this study. The tenets of cognitive theories such as serial--parallel processing, bottom-up--top-down processing, and between-channel redundancy may provide the framework for the manipulation of the subtitling variable in this investigation. Parallel processing of orally and visually cued information can be more effective than the linear processing of the information. Bottom-up processing may act as a back up strategy when top-down processing can not be applied due to lack of background knowledge. Finally, since redundant information presented in two sensory modalities, aural and visual, can be conceived as two separate learning trials, more learning can be expected from audio-print materials.

Research on two educational fields, hearing-impaired and broadcasting, have assessed, with mildly to highly positive results, the effects of subtitles on reading and/or information processing capacity.

As concerns the effects of subtitles in second/foreign language education, early audiolingual research on "text-sound association" already reported positive results showing that simultaneous presentation of the
spoken language and its written equivalent may be beneficial to the
learners. In the same way, research on "reverse subtitling" has proved that
intralingual subtitles (subtitles written in the target language) may improve
comprehension and retention of information in advanced second language
learners. Finally, research on subtitled video has also proven the
effectiveness of subtitles to increase comprehension, retention, and recall of
language input as well as to develop "chunking strategies" which may
release spare language-processing capacity -- especially in intermediate
and advanced learners.

Research on design of learning task provides evidence that
manipulation of the difficulty level of a task affects the learner's mental effort
invested in that task and, subsequently, the learners' performance. In the
same way, there is also evidence that structuring the task content to
"replicate" the organization of a subject's internal knowledge may increase
learning.

Turning to the tests of speaking ability in second/foreign language,
those reviewed here indicate how communicative views of language
teaching practice have equally informed language test development. The
variety of tasks and the performance criteria of those tests reflect the efforts
of the profession at eliciting and measuring the essential features of
communicative language use.
Finally, research on the CALL effectiveness issues collaterally addressed by this study indicates the following: 1) unaided learner control of substantive instructional decisions seems to be less effective than learner control with various forms of guidance or advisement; 2) time on task may provide information about time saving and, more importantly, about learners' strategies; 3) two components of the learning environment, instructional setting and authorship of learning programs, may have an effect on learners' attitudes; 4) word processing appears to improve learners' attitudes toward the writing process, but there is no evidence that it improves the quality of the written product; 5) influenced or not by previous computer literacy, learner's attitudes toward CALL (instruction and subject matter) seem to be positive, although they do not necessarily lead to better performance.
CHAPTER 3
PROCEDURES

Population and Sample

Population

The population from which the sample of this study was drawn was French 2154 and 2155 students at Louisiana State University (LSU). French 2154 and French 2155 are both the fifth in a series of five courses that serve both as an introduction to French and as a fulfillment of the foreign-language requirement for some majors in different LSU Colleges. Both courses are the approximate equivalent of the end of college intermediate-high level study.

The great majority of the population is college-age, with most students being either sophomores or juniors. Almost all of the students enrolled in the two courses have taken the prior semester’s course, French 2102, at LSU.

All sections of French 2154 utilize the same syllabus sequence; the same textbook, Qu’est-ce qui se passe?: Expression orale (Balas & Rice, 72
1990); and similar sets of tests. The course places heavy emphasis upon communicative activities with a considerable amount of time devoted to developing speaking proficiency. Instructors in the three sections of this course utilize authentic materials to stimulate oral activities, and some of them send their students to the language lab to either practice listening comprehension using video or reinforce grammar with drill and practice software.

The emphasis of French 2155 is on reading proficiency and literature. The four sections of the course have the same curricular goals, but each of them utilizes its own literary readings and tests.

Experimental Site

Louisiana State University was chosen as the site of this study for several reasons. First, the spring offerings in French 2154 and 2155 provided an adequate number of subjects to conduct the experiment.

Second, students in both courses had sufficient training in grammar and listening comprehension to permit them to comprehend the basic structures of the language. This allowed the experimental video segments to be "authentic" (contemporary French topics approached by native speakers for native-language speakers), without risking subject abandonment of desired task-level orientation due to linguistic difficulty of those segments. As a result, it was possible to expect that the differences
between experimental treatments reflect the interaction with authentic target language rather than with manipulated language.

Third, the College of Education allowed the researcher free use of its computer laboratory. This meant that the experiment could be conducted with intact classes and little disruption. In addition, the fact that the experiment was held in a location different from the language lab contributed to make subjects unaware that they were participating in an experiment.

Fourth, since the computer laboratory of the College of Education has been conceived for teaching rather than experimental purposes, it was expected that participants in the experiment would have to share it with students from other courses. This expectancy, confirmed later on by reality, was thought to be advantageous: if subjects worked in an environment closer to the real classroom than to the "aseptic" laboratory, experimental results would be more generalizable.

Sample

Sixty students from the three sections of French 2154 and from one of the four sections of French 2155 constituted the experimental sample. The size of the sample was established according to Hinkle, Wiersma, and Jurs's Sample Sizes Table (1988, pp. 666-668). Following this table it was
found that in a two factor design with two levels, using power = .80, \( d \) (standardized size effect) = 0.75, and \( \alpha = .05 \), the required cell size was 15.

The subjects were randomly assigned to the four treatment cells by using a table of random numbers, in accordance with Borg and Gall's drawing procedure (1989, pp. 910-912).

Forty-eight out of the sixty targeted students agreed to participate in the experiment, but four of them declined their participation due to scheduling problems. Thus, the eventual size of the sample, which remained the same during the two experimental sessions, was forty-four students. This remarkably low attrition of the sample was due to the fact that subjects were positively impressed by the courseware during their first practice session, but also and more important, to the fact that recommended strategies (Borg & Gall, 1989) to reduce attrition were applied: 1) emphasis on the importance of the study (subjects were told that the practice sessions were intended to assess the effectiveness of the practice materials); 2) strong commitment of subjects before participation (subjects signed up to participate and chose the days and times of their practice sessions); 3) frequent contacts with subjects in order to maintain interest and rapport (the researcher visited the subjects in their French classes; she invited them to take part in the experiment through
personalized letters (first session) (see Appendix A), and postcards (second session); and, when needed, she contacted them by phone.)

With regard to the subjects' academic background, it was quite heterogeneous as evidenced by the range of undergraduate degrees the subjects were pursuing in the following LSU Colleges: Arts and Sciences, Junior Division, General College, Business Administration, Basic Sciences, Engineering, and Design. Among the subjects from the College of Arts & Sciences, seven were majoring in French.

The computer skills of most of the subjects were limited to word processing, and none of them were familiar with HyperCard software. Five or the subjects were native speakers of languages other than English (Spanish, Japanese, and Vietnamese). And, as concerns sex distribution, the number of females (32) in the sample largely exceeded the number of males (12).

There was no attempt made at identifying the French proficiency level of individual subjects prior to their participation in the experiment.

**Research Design and Treatments**

According to Salomon (1974, 1979), experiments on the instructional utility of media should investigate the ways in which medium's coding elements and related learning tasks interact and affect learning. Referring to
a coding element such as subtitles, Vanderplank (1988) suggests the investigation of their effects on productive skills in second/foreign language. As concerns the learning task, Krendl and Watkins (1983) 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 utilizing a 2 x 2 factorial design. The advantages of this type of design are its efficiency, which allows for the simultaneous study of two variables' main and interaction effects, its control over additional variables which cannot be manipulated (Hinkle, Wiersma, & Jurs, 1988, p. 402), and its precision which prevails over one-way analysis' (Kerlinger, 1973, p. 258). The main drawback of this design is that repeated measures cannot be used, thus many subjects must participate.

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 (subtitles vs. no-subtitles) variable with the oral task level (higher vs. lower) variable:
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.

This 2 X 2 design, illustrated in Figure 3.1, has the advantage of determining not only the main effects of the two independent variables, subtitling and task level, but also the potential of identifying any significant interaction between levels of such variables.

<table>
<thead>
<tr>
<th>Subtitles</th>
<th>No-Subtitles</th>
</tr>
</thead>
<tbody>
<tr>
<td>( A_1 )</td>
<td>( A_2 )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Lower-Level Task</th>
<th>Higher-Level Task</th>
<th>Lower-Level Task</th>
<th>Higher-Level Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>( B_1 )</td>
<td>( B_2 )</td>
<td>( B_1 )</td>
<td>( B_2 )</td>
</tr>
</tbody>
</table>

\( (A_1 B_1, A_2 B_1, A_1 B_2, A_2 B_2) \)

Figure 3.1 2 x 2 Factorial Design: 2 Levels of Subtitling (Subtitles, No-Subtitles); 2 Levels of Task (Lower, Higher)
Variables and Instrumentation

Experimental Variables

Independent Variables

Subtitling. The first independent variable was the media coding element, subtitling. It was a manipulated, active variable with two levels:

1. **Subtitles**: Subjects under treatments 1 and 3 used subtitled video segments to accomplish the required oral tasks.

2. **No-subtitles**: Subjects under treatments 2 and 4 used the same video segments, but without subtitles, to accomplish the same oral tasks.

Task Level. The second independent variable was also a manipulated, active one. It consisted of two oral transactional tasks, description and narration, each of them 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 successfully accomplish the tasks.

1. **Lower-level tasks** were defined as tasks that required the subjects to provide a description or narration which included at least one spatial detail and three specific elements about the people or events they had seen in the video segments.

2. **Higher-level tasks** asked the subjects to provide a description or narration which included various spatial details and at least
five specific elements about the people or events they had seen in the video segments --identical to those of the lower-level task subjects.

It was thought that since higher level-tasks posed more cognitive demands on the viewers than lower-level tasks, 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. But in addition, they had to report on the identity of that 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 those directions.

The two types of tasks, description and narration, were selected for two basic reasons. First, they illustrated the type of transactional tasks which, as mentioned earlier, allow learners to produce "long turns," an important element in the development of communicative abilities. Second, given the subjects' proficiency level, it was assumed that both tasks could be accomplished, at least minimally, by the majority of the subjects.
Indeed, according to the ACTFL Proficiency Guidelines\(^1\), at the
Intermediate-High level of language proficiency "There is emerging evidence
of connected discourse, particularly for simple narration and/or description.
The Intermediate-High speaker can generally be understood even by
interlocutors not accustomed to dealing with speakers at this level...” (1988,
p. 411).

**Dependent Variable**

The dependent variable---*oral communicative performance*\(^2\)---was
defined as achievement in conveying information (describe people, objects,
locations; narrate events, actions) orally.

This variable was measured by applying the instrument developed
by the researcher, a semi-direct test of speaking ability\(^3\), to the transcripts
of the oral samples produced by the subjects.

**Instrument's Development.** 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 achievement at both the
global and component levels of the oral communicative competence
construct devised by the researcher (see Figure 3. 2). At the global level,
"effectiveness" was the major criterion by which subjects were to be
GLOBAL LEVEL

The ability to convey information orally.

Criterion of evaluation

Effectiveness:

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

COMPONENT LEVEL

Grammatical

Knowledge and use of rules governing phonology, morphology/syntax, and vocabulary.

Criterion

Accuracy:

Correctness of the subject's utterances

Discourse

Knowledge and use of rules or devices permitting the structuring of a text with respect to both meaning and form.

Criteria

Organization:

Overall coherence and cohesion of the subject's speech.

Fluency:

Overall smoothness, continuity, and naturalness of the subject's speech.

Figure 3.2 Oral Communicative Competence Construct\(^{(n)}\)
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 B, Oral Communicative
Performance: Evaluation Form) designed according to the following models:
1) Bartz's (1979) Amount of Communication and Accuracy scales were used
to build the Effectiveness and Accuracy measures; 2) O'Malley, Chamot,
Stewner-Manzanares, Russo, and Küpper's (1985) speech organization
criteria were taken into account to build the Organization scale; 3) Emmett's
(1985) Fluency scale provided the basis for the Fluency measure; 4) Simon's
(1986) communicative measurement procedure inspired the form of the
evaluating instrument.

Since "effectiveness" was considered as the primary criterion of oral
communicative performance, the points gained in the Effectiveness scale
were weighed by three and added to the points obtained on the remaining
scales. Incorrect or inaccurate information did not count against the scores
obtained on the different scales.

Instrument's Validation and Application. A native speaker, college-
level French instructor was asked to evaluate the subjects' oral samples. To
proceed with the evaluation, and before the practice sessions, the researcher
showed the evaluator the courseware, "Practicing Spoken French" (PSF), which subjects were going to use in such sessions. Namely, she showed her how to enter a "PSF User" stack, how, once within a stack, to reach a subject's oral sample from the "Control Card," and how to control the videodisc player from the computer screen. The researcher then asked the evaluator to watch the video segments several times to become thoroughly acquainted with them.

After conclusion of the practice sessions, the researcher provided the evaluator with photocopies of all subjects' evaluation forms, and she trained her on how to use them. The forms already contained the researcher's non-rated transcripts of the oral samples recorded by the subjects during the two practice sessions (see sample transcripts, Appendix C).

In order to insure inter-rater reliability, ten oral samples were randomly selected from the first experimental session\(^{5}\). Both the evaluator, who was unaware of the experimental treatments, and the experimenter listened to the selected samples and separately scored the corresponding transcripts. The inter-rater reliability coefficients of the oral performance instrument established after this procedure were: \( r = .96 \) (Overall Oral Performance); \( r = .83 \) (Effectiveness); \( r = .95 \) (Accuracy); \( r = .94 \) (Organization); \( r = .91 \) (Fluency). Those coefficients being highly positive, the instructor proceeded to the evaluation of the remaining oral samples: she went through every
subject's stack, listened to each recorded sample, verified each of the
transcripts, and recorded the corresponding scores on each of the evaluation
forms copies. Overall, it took about fifty-four hours to validate and apply the
oral performance instrument.

Non-Experimental Variables

Attitude Questionnaires

Although not an experimental variable, a Likert-type measure,
consisting of two attitude questionnaires (see Appendices D and E), was used
to provide descriptive data about the subjects' attitudes toward their speaking
experience with multimedia courseware, and, particularly, with "Practicing
Spoken French."

The two questionnaires, forms A and B, were given at the end of the
second practice session. Questionnaire form A was administered to subjects
in the subtitles condition (T1, T3) while questionnaire form B was
administered to the remaining subjects (T2, T4).

The questionnaires consisted of 10 statements to which subjects were
asked to respond by choosing between: (1) strongly disagree; (2) disagree;
(3) no opinion; (4) agree; (5) strongly agree. Since it was assumed that
subjects were unaware that they had practiced under a variety of treatments,
the statements were written in a manner so as to maintain such unawareness.
Statements 1-3, identical for both questionnaires, addressed subjects' attitude

toward their practice with multimedia courseware. Statements 4-10 targeted subjects' opinion about particular aspects of "Practicing Spoken French" design; these statements were also identical for both questionnaires except those asking for the subjects' opinion about subtitles (statements 8 & 9, questionnaire form A) or videodisc features (statements 8 & 9, questionnaire form B).

Prior to being utilized in the pilot study, each statement of the questionnaires was examined in light of the informal criteria suggested by Edwards (1957, pp. 13-14) for editing statements to be used in the construction of attitudes scales. These criteria include avoiding the use, among others, of the following: difficult words, double negatives, factual or misleading statements, and statements containing universals ("only," "always," "never," etc.) which may introduce ambiguity.

Likert's (1932) criterion of internal consistency, which is based upon the difference between the average score of the most and least favorable individuals, was calculated on the answers of the pilot study's subjects to questionnaire form B\(^{(0)}\) to determine the index of discrimination of each statement. Following Likert's design, the means of the individual statements' responses from the top 20% of the subjects (most favorable) were compared to those from the bottom 20% (least favorable). It was determined that, on a five-point scale, a statement with a discrimination index of 1.0 or above would
qualify for inclusion on this descriptive instrument. All comparisons resulted in positive indices, which indicated correct directional interpretations of statement wording. The statement which produced the lowest discrimination index (1.0) was worded, "More practice sessions similar to the ones I had would increase my confidence to talk to other people in French." The statement that was retained which produced the highest discrimination index (3.0) was worded, "The time and the number of tries allowed to record the description or the narration were sufficient."

Procedures and Data Collection

This experiment took place at the LSU College of Education's computer laboratory from April 13 to April 29, 1992. To minimize the Hawthorne Effect, that is to avoid the fact that subjects' awareness that they were participating in an experiment might distort the results, no mention of the variety of treatments was made. Subjects were told that the College of Education had invited them to participate in two French practice sessions using computers and interactive video. They were informed that the College wanted to know their reactions to the "Practicing Spoken French" (PSF) courseware in order to assess its effectiveness and make decisions about its implementation for French 2154 and 2155 students in the future.
Two weeks before the experiment, instructors of the experimental French sections were offered a demo of "Practicing Spoken French," which impressed them favorably. The instructors were instrumental in keeping the sample's attrition rate low: they encouraged their students to participate in the practice sessions, and two of them excused their classes from attendance at regular class times.

One week before the experiment, participants were told to sign up on a scheduling grid the days and times they preferred to do the practice sessions. Since those sessions had been estimated to last ninety minutes each, they were scheduled throughout the days of the experiment at intervals of two hours. Participants were also told to meet for those sessions in room 118 of Peabody Hall on the Louisiana State University campus.

Four experimental workstations, consisting basically of a computer, a TV monitor, and a videodisc player, were set up in the computer lab. Each day of the experiment, and prior to the arrival of the participants, the researcher loaded the computers at the four workstations with as many copies of the "PSF User" stack, which featured the tasks users could perform in their practice sessions, as subjects had signed up to practice during that day. The researcher then "prepared" each of those copies by entering the name of the subject who was going to use it, and selecting the type of treatment to which the subject had been assigned.
Also each day, after the departure of the subjects, the researcher compressed, copied to a diskette, and removed from the computer the stacks of the subjects who had done their practice that day. This procedure was aimed at keeping enough disk space available for the stacks which had to be loaded the following day.

As the subjects arrived at the computer lab on the days and times they had selected for their two practice sessions, they were told to sit at one of the workstations and enter the stack bearing their name. Subjects were told to make use of the earphones available at each workstation. The earphones, connected with the computer and the TV monitor via a Y stereo adapter, were intended to reduce disturbing sounds and increase subjects' concentration on the experimental materials.

Also available at each workstation was information about how to enter and exit the "PSF User" stack, and how to get French characters on a Macintosh U.S. keyboard (see Appendix F).

Once within the "PSF User" stack of the "Practicing Spoken French" courseware, subjects worked on the unit of their choice ("Description" or "Narration") of either Program 1 (1st experimental session) or Program 2 (2nd experimental session). Each unit was comprised of four working stages which 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 subtitled or unsubtitled 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 read in the video segment.

3. **Drafting one's own oral sample.** Subjects were given formal guidance to prepare an oral sample based on the segment they had seen. They were told to write a draft of a description, or a narration, following the higher/lower requirements of the task corresponding to the treatment to which they had been ascribed. Throughout this stage, users were able to watch the video again or get linguistic help from the "Très Utile" stack.

4. **Recording oral sample.** Subjects recorded the samples they had previously drafted. The allowed number of recordings was limited to five and the length of each recording to three minutes.

Taking into account cognitive psychologists' advice (Wittrock, 1983; McDonough, 1981) about meaningful learning, stages 1 to 3 were aimed at
insuring that subjects were adequately prepared prior to undertaking the last and most important of the stages, speaking. It was thought that the activities of these three stages would help subjects to improve their comprehension of the video segments and their memory of the people and events in those segments.

Moreover, it was expected that the four stages would give subjects the opportunity to apply some sets of cognitive strategies which according to language specialists (Oxford, 1990) may provide a rich and powerful support to any language learning effort. Namely, analyzing (analyzing expressions, translating), creating structure for input and output (taking notes, summarizing), practicing (repeating, formally practicing with sounds and writing systems, recognizing and using formulas and patterns, recombining).

Subjects were encouraged to accomplish the stages as well as possible. And, even if the first of the two experimental sessions was conceived by the researcher as a training one, nothing was said to the subjects in order to keep their performance's standards equally high in both sessions.

During those sessions, unobtrusive descriptive data were taken by the computer as well as by the researcher. The computer data included the total and detailed time spent on task, the number of times subjects entered the "Très Utile" stack, the number of times they recorded their oral samples, and
the length of those samples. The researcher data was based on her systematic observation of the subjects' practice and oral samples' drafts and transcripts. Additional descriptive data were provided by two Likert-type questionnaires completed by participants at the end of their second practice session. Computer, researcher, and Likert data were used to better understand and interpret the effects of the independent variables of this study; the data were also used to provide additional insight on some CALL research issues.

After implementation of the experiment, which extended over fourteen days and took about one-hundred and twenty hours, subjects were thanked for their participation and were given an explanation about the nature of such participation. They were told that they had collaborated in an experiment, that simultaneous measures had been taken on their speaking practice sessions, and that they had practiced under different treatments.

**Pilot Study**

A pilot study using four subjects from one section of French 2155 and six subjects from one section of French 4100 was conducted at Louisiana State University on April 9 and 10, 1992. The purpose of the pilot study was: 1) to verify the feasibility of the experimental procedure; 2) to determine the average time it would take the experimental subjects to accomplish the PSF's
units, 3) to test for potential instructional/technical flaws in "Practicing Spoken French," and 4) to assess the internal consistency of the attitude questionnaires.

Six subjects were assigned to two (T1 and T3) of the four treatments in this study, and four subjects were assigned to the remaining ones (T2 and T4). From their practice, subjects in the four treatments showed that the devised procedure could be implemented. All of them went through the four stages of at least one instructional unit in the way it was expected and with none, or few, technical problems. Six of the subjects typed their oral sample draft(s), and all recorded them at least twice.

Subjects who participated on the first day of the pilot study were told to work on the two units of the first PSF program. None of them could complete the units within the approximately 120 minutes that the researcher had allocated for such units. Therefore, subjects who came on the second day were told to work on only one of the units -- which took them 70 to 90 minutes to complete. In light of this finding, it was decided that, in both practice sessions, experimental subjects would work on only one instructional unit, and it was estimated that they could complete it within a ninety-minute period.

Four minor changes were made to "Practicing Spoken French" as a result of the pilot study. First, the text of the help window of the "Home Card" was modified to let experimental subjects know that they had to work only in
one unit—instead of two. Second, the text of the help window of the last card of the "Très Utile" stack, was also modified because it was observed that pilot subjects did not understand the directions for getting the sound of the words displayed on the screen. Third, based on the systematic observation of the subjects as they were working on the units, changes on the PSF background script were made so that all the help windows of the "PSF User" stack were closed the first time the subjects entered it. Indeed, displaying the help windows opened the first time the subjects accessed the stack seemed to make them resent those windows as "imposed" lengthy pieces of information. It was then thought that closing the windows would avoid this impression and increase subjects' feeling of "discovery" and control of the program. Fourth, the script of the "A quiz" card of both PSF programs was also modified to avoid invisible characters (e.g., "Tab" or "HRT") typed by the users before entering the required answer on the provided spaces, making the computer scoring inaccurate.

Finally, Likert's criterion of internal consistency was computed on the pilot subjects' responses to attitude questionnaire form B. The criterion showed the questionnaire's ability to discriminate between positive and negative reactions to speaking practice with multimedia courseware.
Data Analysis

Scores of oral samples recorded by the experimental subjects in the two experimental sessions were submitted to two-way factorial analyses of variance (ANOVA), using the SuperANOVA (1991) software package on a Macintosh Ilsi. 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").

Finally, the Statistical Package for the Social Sciences, SPSS\textsuperscript{x} (1986), was also utilized, on the LSU Computer Network Computer Center, 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 observations 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.

**Summary**

This study investigated the effects of subtitled vs. unsubtitled video during transactional task practice with multimedia courseware upon oral communicative performance of college students of French. In order to appreciate more fully and determine the context of any effects, one additional independent variable was considered, level of oral transactional tasks (higher-level vs. lower-level). This 2 x 2 factorial design utilized a composite oral performance score, assigned to the oral samples recorded by the experimental subjects, as the dependent measure.

Furthermore, computer and researcher's observations and two attitude questionnaires were utilized to provide useful descriptive data from the study. These results were not considered part of the research design but they served to better understand the experimental findings, to provide ideas about how to improve the courseware, to generate hypotheses for future studies in the area, to increase knowledge on CALL effectiveness and, ultimately, to prove that quantitative and qualitative research paradigms can be successfully used in combination.
CHAPTER 4
COURSEWARE DEVELOPMENT

Background to PSF

"Practicing Spoken French" (PSF) was created for use in the doctoral research project described in this dissertation. Courseware planning and development was undertaken by this researcher and one computer programmer in August 1991. In April 1992, after having worked on it for about 715 hours (see Appendix G), PSF was finished and ready for experimental use.

Software and Hardware Selection

Two software tools, HyperCard 2.0 (1987) and Voyager Videodisc 2.2 (1987) were used in the development of the courseware package. While the first of these tools provided the "physical" structure of the courseware lessons, the second one allowed for the control of the videotape segments by the computer program.

The courseware 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:
. A Macintosh Ilsi microcomputer with 5mb RAM and 80mb Internal Hard Disk.
. An AppleColor High-Resolution RGB monitor (13").
. A Pioneer videogisc player (LD-V2200).
. An Amdex Color 300T TV monitor.

The first of the above components, the Macintosh Ilsi, was chosen because its built-in recording capabilities made this computer particularly appropriate to handle one of the objectives of the experiment which was to have subjects record their own oral samples with ease. The last of the components, the printer, was included in the workstation for something more than having users get a copy of their oral sample’s drafts. The researcher thought, and results ratified her guess, that allowing users to “touch” their typed “products” could increase users’ willingness to work harder in the preparation of the oral tasks.

Courseware Configuration: Rationale

The courseware integrates both interactive videogisc (IVD) and hypermedia, two technologies whose instructional potential can no longer be denied. With respect to IVD, there exists a significant amount of discussion (McCoy & Weible, 1982; Bosco, 1984; Clark, 1984; Smith, 1987,
Goforth, 1992) concerning the instructional features of the medium, namely about IVD's potential to increase user's interaction with large amounts of instantly accessed video-based materials. Within the interactionist perspective of Second Language Acquisition (SLA), the potential of IVD for language learning has also been extensively highlighted. As Daughty (1991), for instance, observes:

If language learning is seen as a regular and ongoing interaction between the learner's mental abilities and the linguistic environment, each contributing to language acquisition and each influencing the other, then 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)

With regard to hypermedia, theoretical discussions (Conklin, 1987; Morariu, 1988; Megarry, 1988; Jonassen, 1988b, 1988c, 1989), and the implementation of sophisticated prototypes (Nakajima, 1988; Underwood, 1989; Ashworth & Stelovsky, 1989) seem to support the significance of the software as far as the creation of instructional materials is concerned. Indeed, with hypermedia the human involvement is clearly central: it provides an "enabling" rather than a "directive" environment which requires users to constantly make decisions, thus forcing them to apply higher order thinking skills. Also because of its knowledge representation format, hypermedia can empower instructional design based on cognitive principles.
such as the structured hypertext model (Jonassen, 1986) and its underlying theoretical basis, the generative learning processing hypothesis (Wittrock, 1974).

**Courseware Design**

Wittrock (1974) contends that meaning from material presented by computer or any other medium is generated by activating and altering existing knowledge structures in order to interpret what is presented. The amended structures are then encoded in memory as distinctive features that may be accessed later to explain or interpret new information. In the same way, Jonassen (1988a, p. 155) observes that, "The emphasis of instructional activities embedded in courseware should be on facilitating knowledge acquisition by making the learner processing more generative." Subscribing to these views, "Practicing Spoken French" (PSF) features, as shown in Figure 4.1, a menu-driven design which: 1) allows users' free choice of tasks and modules, 2) indicates users' completion of modules within a task, and 3) provides users with convenient points for exiting and reentering the modules.

**How PSF Looks**

Building upon a variety of audio-visual resources which include animations, graphics, music, recorded voices, and videodisc segments (see
Figure 4.1 "Practicing Spoken French": Menu-Driven Design.
Appendix H), "Practicing Spoken French" is made up of three sequences of HyperCard cards (or screens) called stacks: 1) the "PSF Resources" stack which contains pictorial and textual information; 2) the "PSF User" stack which has 57 cards featuring the tasks users perform with the courseware; and 3) the "Très Utile" stack which has 22 cards containing specific information about the tasks, and which can be accessed from a particular card of the "PSF User" stack.

The "PSF User" stack consists of two similar programs, each of them having two distinct units, "Description" and "Narration," organized around two modular extensions. An overview of the "Description" Unit of Program 1, illustrated by Figures 4.2 to 4.15 (see pp. 105-118), follows.

**Introductory Screens**

To get into Program 1, users double-click on the stack of their name. Once within it, they go through seven introductory screens which contain information about the objectives and prerequisites of the program. One of the screens, displays a picture of various faculty members at LSU and asks users to click on one of the members' name to listen to what he says about the program (see Figure 4.2). Another screen, "How to Move Around," shows and explains the meaning of the icon-buttons which users will encounter throughout the program (see Figure 4.3). A final introductory
screen, the "Home Card," shows users how to get into the various modules of the two units of the program (see Figure 4.4).

Module 1

If users decide to visit the first module of the "Description" unit, they click on the Module 1 button of that unit, and a screen consisting of a scanned picture (BR University Lake) and a caption (definition of description) appears (see Figure 4.5). Once within the module, users enter the "What to See" screen to view the video segment they will have to describe later. To reach the video segment, users click on a TV button at the top right of that screen and a "palette" with the control buttons of a videotdisc player appears. After viewing the video segment, users may check their comprehension by attempting to answer six questions of a trivial pursuit game for which scoring and feedback procedures have been properly ensured (see Figures 4.6, 4.7, & 4.8).

Module 2

The second module of the "Description" unit begins with a screen featuring a picture and a caption similar to those of the initial screen of the first module. Next, there is the "What to Describe" screen which asks users to depict the main character of the video segment seen on the previous module. Later, in the "Your Draft" screen, users may type and print what they are going to describe orally (see Figure 4.9) following the directions...
found on a Help window (see Figure 4.10). From the "Your Draft" screen, users may call upon the video segment or enter the "Très Utile" stack to explore its three sections: 1) grammar; 2) learning how to describe; and 3) difficult words (see Figure 4.11). The last of these sections provides for the translation, listening, and practice of ten words heard on the video segment (see Figure 4.12). The next to last screen of the module introduces the "Recording Studio." There, users record the descriptions they have drafted following the recording procedure explained on the Help window (see Figure 4.13). On the last screen of the module, "What to Do Next," users may decide between leaving the program or continue working on the "Narration" unit.

If users opt for the second decision, they enter the first module of "Narration" and a flow of screens similar to the ones described above follows.

A "Control Card" (see Figures 4.14 and 4.15) orchestrates the three "Practicing Spoken French" stacks. This card, not accessible to the user, allows the instructor: a) to enter the user's personal data; b) to select the program number (1 or 2), the task level (higher or lower), and the video segment type (subtitled or unsubtitled); c) to define the videodisc settings; d) to access the information recorded by the computer about the user's
Figure 4.2  PSF screen showing the user how to get the introductory message uttered by one of the people portrayed on the scanned picture.
HOW TO MOVE AROUND

Look at the "buttons" (icons) at the top and the bottom of the screen which you will see as you navigate through this program. If you want to know what these buttons allow you to do, put the pointer (mouse) over them and information will appear automatically.

When you are ready to continue, click anywhere in this card.

Figure 4. 3. "How to Move Around" screen showing a range of icon buttons used in PSF.
Figure 4.4 PSF "Home Card" screen. This screen allows the user to enter/exit the various modules of the program.
Description:

"A verbal account or portrayal
of a person, scene, etc."

(The New Lexicon Webster's Dictionary, 1989)

Figure 4.5 PSF introductory screen of the first Module of the "Description" unit. The scenery depicted in the screen is meant to enhance the program with a "local" flavor.
TRIVIAL PURSUIT
(Question 1)

What is the name of the young French advertising "star" portrayed in this video segment?
(You may either click on the answer of your choice, or type it into the box at the lower right.)

Gregory Marti  Maxime
Luc Châtelain  Yves Saint-Laurent

Your Answer: Luc Châtelain

Figure 4.6  Screen showing a user's answer to one of the questions in the PSF trivial game. The user may have either typed the answer in full or popped it up into the answering box by clicking in one of the four possible answers.
TRIVIAL PURSUIT

The trivial pursuit is finished!

Your score is 6 out of 6.

BRAVO! Joe

You have fulfilled our expectations.

The key to your answers?... Next screen.

(11 of 13)

Figure 4.7 User’s scores in the PSF trivial game. The scoring procedure includes two rewarding strategies consisting of: 1) three personalized messages (for "excellent," "good," and "poor" results); and 2) a "stimulating" musical excerpt (for "excellent," and "good" results).
The box below shows your answer, and the correct one, to each of the questions of the Trivial game. Use the arrows in the box to move between questions.

If you want to view the video again to check why your answers are, or aren't, correct, click on the TV button.

**Question #1**

**Question:**
What is the name of the young French advertising "star" portrayed in this video segment?

**Correct Answer:**
Luc Châtelain

**Your Answer:**
Luc Châtelain

**Points:**
1

Figure 4.8 Feedback screen which shows both the correct answer and a user's answer to one of questions of the trivial game. The screen allows users to go back to the video segment to self-check their answers.
Figure 4.9 The "Your Draft" screen. In this screen users have to draft a description about the main character in the video segment seen during the first module of the program. Pressing the bulb button forwards users to the "Très Utile" stack; pressing the TV button takes them to the source of their description.
Using the computer keyboard, or pencil and paper, draft a description, in French, of Luc Châtelain, the young French advertising "star" portrayed in the video segment "Les enfants et la publicité".

Your description has to include the following:
1) one of the two settings where Luc appears inside his house (eg: kitchen, study room)
2) Luc's physical appearance (estimated age, body complexion, hair, face, etc.)
3) Luc's clothes

* * * * * *

Click on the BULB button at the top right of the screen to get some ideas about how to organize your description, and to see and practice some useful vocabulary.
If you feel that you need to see the images again, click on the TV button.
If you need more space to type, click on the up/down arrows.
Whenever you feel ready to record your description, stop writing, print your draft if you have typed it, and move on to the next screen.
Click on the ? button each time you want to see/hide this information.

Figure 4. 10 The "Your Draft" Help window. The window shows users which directions to follow to prepare their descriptions.
You have reached the “Très Utile” stack. Here, you will find information which can help you to better prepare and deliver your description. That information deals with:

I. Grammar
II. Learning how to describe
III. Difficult words

Click anywhere to continue.

Figure 4.11 The “Très Utile” stack: title page screen. Reached from the “Your Draft” screen, the stack contains information which may help users to improve their descriptions.
III. Difficult words.
Click on the 7 button to learn more about the words and expressions below that you have heard in the video segment:

1. Prise de vue
2. Joues
3. Taches de rousseur
4. Avoir les cheveux en brosse
5. Marrant
6. Taille
7. Roux
8. Dos yeux grands ouverts
9. Lèvres
10. Avoir l'air

Avoir les cheveux en brosse (expression figurée): to have a crew-cut.

Figure 4.12 The "Difficult words" screen of the "Très Utile" stack provides the translation, and listening and repetition practice of ten words uttered during the video segment.
You have entered the "Recording Studio". Click on the ? button to see the directions you must follow to record your speech.

Good Luck!

When the message appears asking you to click on "OK" or "Cancel", be sure to ALWAYS click on "OK".

Figure 4.13 The "Recording Studio" screen. This screen allows users to record, up to five times, the description previously drafted on the "Your Draft" screen.
Figure 4.14 The PSF "Control Card" allows for: 1) the selection of program number, task level, video segment type, and videodisc setting; 2) the recording of the user's personal data, spent time, and performance; 3) the access to the user's drafted and recorded oral sample.
Figure 4.15 "Control Card" screen displaying the Quiz score and the time spent by a PSF user in the "Narration" unit of Program 2.
performance and time spent on the various parts of the program; and e) to see and listen to the user's drafted and recorded oral sample.

**How PSF Works**

If the screens mentioned above provide a detailed picture of the courseware, the flowchart below (see Figure 4.16) gives an idea of the courseware's assembling and how users may maneuver within it. Stages where it is possible to choose between two or more options are indicated by diamond shaped boxes. Stages where video segments can be accessed are shown with rounded rectangle shaped boxes. When additional displays are available from a particular screen, they are marked with small rectangles fitted into bigger ones. Initial and final decisions are signaled by circles.

By following the flowchart, it is also possible to uncover the "events of instruction" embedded into the courseware design. These "events," which according to Gagné and Briggs (1979) may be used to influence the internal learning processes, are sequenced as follows:

1. Gaining attention (screens 1, 2, 3)
2. Informing learner of unit objectives (screens 4, 5, 9, 11, 22)
3. Stimulating recall of prior knowledge (screens "Très Utile" 1-6)
4. Presenting stimuli with distinctive features (screens 3, 5, 8, 21, 25).
5. Providing learning guidance (screens 6, 7, 10, 20, 22, 23, 27, and screens "Très Utile")
Figure 4.16 Flowchart Program 1. Description Unit.

- Screen 27 (What to Do Next)
- Screen 26 (Recording studio)
- Screen 25 (Ready to record)
- Screen 24 (Your yarn)
- Screen 23 (Mending for)
- Screen 22 (Reports)

- Yes
- No

- Continue to Narrator Module 1 or another

- See "This Unit"

- Continue

- Grammars
- - How to describe
- - Difficult words
  (see, listen, repeat)
6. Enhancing retention (screens 12-17, 19, 24)
7. Eliciting performance (screens 12-17, 24, 26)
8. Assessing performance (screen 18)
9. Providing feedback (screen 19)

Unfortunately, due to the inability of technology to process and respond to the complexities of natural language in a way that effectively simulates the human brain, events 8 and 9, assessing performance and providing feedback, have been applied only to the user's comprehension of the audio-visual materials.

**Technical Features**

Using the design guidelines described earlier, "Practicing Spoken French" is programmed to offer the user and/or the instructor the following possibilities:

1. Run and play back a video segment (with or without subtitles).
2. Listen to text displayed on the screen.
3. Run and play back a native speaker/instructor's audio message.
4. Run and play back the user's voice.
5. Go to the beginning/end, slow down, freeze, fast forward, and rewind videodisc images.
6. Show or hide help windows.
7. Go backward or forward to other screens or stacks.
8. Scroll through text or a section of text.
9. Record subjects' scores on listening comprehension, and provide immediate feedback on correct and incorrect answers.
10. Keep track of the time spent by users on the different stages of a program unit (viewing, answering video-related questions, drafting and recording an oral sample).

Selection and Production of the Experimental Video Segments

Selection

The visual input for "Practicing Spoken French" was provided by four video segments selected from France Panorama (1991), a quarterly video magazine containing contemporary "authentic" materials\(^5\) from France, and designed for use in school and university French second-language programs. The paragraphs that follow describe the process used for selecting the experimental video segments.

First, sixteen segments were preselected by this researcher from over 10 hours of the available video source (see Appendix I). The segments were between two and four minutes in length, reflecting the recommendations from prior research (Lavery, 1984; Garza, 1986) on optimal segment length of video materials intended for foreign language teaching.
Next, the preselected segments were submitted for quality and difficulty evaluation to four instructors who were teaching, or had recently taught, French 2154 at LSU. The instructors were asked to rate the preselected segments on three "quality" criteria (situational appropriateness, inherent interest value, and audio/visual correlation), and one "difficulty" criterion (grammatical and lexical complexity) (see rating instrument, Appendix J).

Finally, following the instructors' ratings (see Table 4.1), four segments were retained for use by subjects under the four treatments during the two sessions of the experiment. The overall quality and difficulty mean ratings for the segments of the first session were respectively 4.08 and 2.87 (1 on the quality scale was "very poor" and 5 was "excellent"; 1 on the difficulty scale was "very easy" and 5 was "very difficult"). The quality and difficulty mean ratings for the segments of the second session were 4.12 and 3.00 (see selected video segments, Appendix K).

Production

After selection, the four video segments were produced in collaboration with various individuals and companies (see Appendix L). The production followed these stages: 1) recording; 2) subtitling; and 3) videodisc mastering.
Table 4.1 Mean Ratings of Preselected Video Segments on Quality and Difficulty Criteria.

<table>
<thead>
<tr>
<th>Quality criteria</th>
<th>Difficulty criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Situational appropriateness</td>
<td>Grammatical &amp; lexical complexity</td>
</tr>
<tr>
<td>mean</td>
<td>mean</td>
</tr>
<tr>
<td>Segment 1</td>
<td>4.5</td>
</tr>
<tr>
<td>Segment 2</td>
<td>4.0</td>
</tr>
<tr>
<td>Segment 3</td>
<td>2.74</td>
</tr>
<tr>
<td>Segment 4</td>
<td>4.0</td>
</tr>
<tr>
<td>Segment 5</td>
<td>4.0</td>
</tr>
<tr>
<td>Segment 6</td>
<td>3.75</td>
</tr>
<tr>
<td>Segment 7</td>
<td>4.25</td>
</tr>
<tr>
<td>Segment 8</td>
<td>4.0</td>
</tr>
<tr>
<td>Segment 9</td>
<td>4.0</td>
</tr>
<tr>
<td>Segment 10</td>
<td>4.0</td>
</tr>
<tr>
<td>Segment 11</td>
<td>3.0</td>
</tr>
<tr>
<td>Segment 12</td>
<td>4.5</td>
</tr>
<tr>
<td>Segment 13</td>
<td>3.0</td>
</tr>
<tr>
<td>Segment 14</td>
<td>4.0</td>
</tr>
<tr>
<td>Segment 15</td>
<td>4.5</td>
</tr>
<tr>
<td>Segment 16</td>
<td>4.25</td>
</tr>
</tbody>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
**Recording**

The selected video segments were initially recorded on a SMPTE 1" tape. The four segments\(^{(7)}\) were recorded a first time, and after insertion of a video black time period of 20 seconds, they were recorded again. Each of the segments was externally edited by addition of a preceding light blue screen displaying the title of the segment during 4 seconds. Segments' time codes in and out were opportunely tracked.

From the initial SMPTE 1" tape, segments were transferred to SVHS because the equipment used to insert the subtitles required that format. The time codes of the 1" tape were put in the audio channel of the SVHS tape.

**Subtitling**

Four out of the eight segments included on the master tape were overlaid with subtitles which appeared at a rate of three lines per screen and remained visible for about 3 1/2 seconds. The subtitles included the verbatim\(^{(9)}\) of the video segments being tested, and they were written in two fonts: 95 Helvetica (words of the narrators and interviewers), and 96 Helvetica Oblique (words of the people interviewed).

The equipment used to generate the subtitles consisted of a Macintosh IIci, a NuVista+ card, a DisQuest card, a read/write optical video drive, a SVHS drive, and numerous monitors and stereo equipment. The DisQuest card was used to read the timecode from the SVHS deck and to
control the operation of the deck. Furthermore, the DisQuest card generated external synchronization signals, which were necessary to record from SVHS to optical drive.

The overlaying was then done live. The SVHS deck generated the original video, the NuVista card overlayed the video image onto the Macintosh graphics generated by a custom-built software system, and the optical drive, using the external synchronization generated by the DisQuest card, recorded the information.

After the overlaying, the optical drive was directly connected to the SVHS and the video was dubbed onto the latter portion of the SVHS tape so that the tape could be ready for videodisc mastering.

Videodisc Mastering

Eight video segments, four subtitled and four unsubtitled, were then mastered on one sided C.A.V. (Constant Angle Velocity) plastic disc with a maximum running time of 30:00. An Optical Disc Corporation's (OCD) 610-A recorder was used in the transferring process from the videotape to LaserVision videodisc. The OCD system was adopted because of its two major advantages: 1) its ability to transfer from different tape sources (1", 3/4", Betacam, Betacam SP or D2), and 2) its fast turn around (usually within 2 work days) without going through the lengthy process of manufacturing molds which can cost thousands of dollars for the smallest runs.
Using PSF: Optimizing Comprehensible Output

As mentioned earlier, learning to talk in "long turns" seems to be essential for the development of communicative skills. However, this condition being important, it is not sufficient; ability to communicate has to be measured in terms of speech's quantity and quality. As Swain (1985) and Gass (1988) observe, in contrast to Krashen (1982), comprehensible input is not enough to promote second language; learners must also engage in comprehensible output in order for the morphosyntax of the interlanguage to develop sufficiently to reflect a level of mastery of the target language. Agreeing with this view, the main goal of "Practicing Spoken French" has been on getting users to produce oral samples whose length and, particularly, comprehensibility might be judged as "optimal" by French intermediate level standards.

Two characteristics of the courseware, a tension-free environment and the systematization of the design, contribute to attaining this goal. First, attempts are made to avoid giving users the feeling that they are coerced, or pressured by the computer: a) by adding pictures of local people or landscapes, which gives PSF users the feeling that the package is specially designed for them; b) by including some relaxing as well as some rewarding music excerpts; c) by writing the directions of the program in a friendly tone; and d) by maximizing the number of choices available to users within the
package (e.g. type of task, number of times to see the video segments, number of times to enter "Très Utile," or number of times to record the oral sample). By virtue of these four factors, PSF may release users from their reluctance to speak, record, and listen to their own voices in the target language, thus increasing their confidence to participate later on in real communicative situations.

The second characteristic, the systematization of the design, may have specially contributed in building up the so-called "performance memory," (Rivers, 1991) and, in turn, increasing the accuracy of the oral output. Such systematization is characterized by: 1) sequential ordering of the instructional events discussed earlier; 2) segmentation of the programs' tasks into units and modules which allows the practice of the four language skills in two progressive sequences- Reading-Listening-Writing (Module 1), and Reading-Listening-Writing-Speaking (Module 2); 3) hierarchical organization, in a systematic top-down fashion, of the information presented in "Très Utile" (Grammar, How to Describe/Narrate, and Difficult Words), which allows learners to put into practice, and to integrate into their semantic networks, grammar rules and lexical input.

This emphasis on systematization is both theoretically and empirically grounded. Indeed, cognitive psychology theorists (Bruner, 1977; Ausubel, 1968; Rumelhart, D. & Norman, D., 1978; Anderson, 1980; Karmiloff-Smith,
1986) have noted that learning requires more than gaining automaticity. According to these theorists, learning is the process by which organizational structures (e.g., schemata) are constantly modified in accordance with and for the purpose of integrating new information. And, as Anderson states, "if research on human memory has shown one thing, it is that the organization of material is of central importance." (p. 16)

Sharing this viewpoint, second language acquisition theorists suggest that learners should be provided with the organization that enables them "... to see where they are going," (Rivers, 1991: 5), or that "the learner needs to impose organization and to structure the information that has been acquired." (McLaughlin, 1987: 136). Also pleading in favor of systematization, instructional designers observe that freedom of choice, this key facet of hypermedia, may turn out to be a main drawback; and, as Marchionini (1988) puts it, "... a rich learning environment can become an environment of 'hyperchaos'... freedom to learn is not a sufficient condition to assure effective learning." (p. 10).

Underscoring the above views, a number of experiments on instructional design (Shavelson, 1972; Reif & Heller 1982; Eylon & Reif, 1984) have shown, as mentioned in Chapter 2, the correlation between the organization of learning materials, and the resulting organization of internalized knowledge.
Conclusion

This chapter has had at its major focus to describe the development of the multimedia package "Practicing Spoken French" (PSF) which could illustrate a successful use of computer and videodisc technologies for teaching French speaking skills. More precisely, the chapter has looked at the package's background, authoring tools and hardware, configuration, design, and central emphasis.

Although results of the experiment conducted with PSF, were highly positive, the courseware should not be perceived as the panacea of pedagogical resources for speaking practice. PSF has limitations. First, as any educational software, the effectiveness of PSF is somewhat limited to the rule-governing manipulation of formal structures (Rivers, 1991; Winograd & Flores, 1986); and because of this limitation, practice with PSF can prepare, but not replace, those kinds of speaking practice which involve human interaction. Second, this is one of the first projects of its kind, and there are certainly many improvements to be made, such as the computer's assessment and feedback of the oral tasks which will only be solved as technology improves. Third, PSF is in need of further testing in different environments and with greater number of subjects.

In spite of these limitations, PSF has made significant strides in creating a learning environment which is both "pleasurable" and "sound." Its
pleasurability is largely evidenced by the users' positive reactions. And this is certainly one of its main assets since learning, as sensitive educationalists remind us (Doll, 1989), should be an enjoyable experience. Its soundness is evidenced by the powers which it bestows upon users. Without dismissing arguments which rightfully criticize the mediocrity of some marketed software (Roszak, 1986), "Practicing Spoken French" gives users the power to: a) process audio-visual information at various depth and length levels, b) prepare two kinds of communicative tasks according to a suggested framework, and c) bridge the gap between the written and the spoken word. But maybe the best evidence of the courseware's soundness is provided by one student who liked PSF because it gave him "... a feeling of achievement."
CHAPTER 5

RESULTS AND DISCUSSION

Introduction

This study investigated the effects of subtitled video segments during transactional task practice on oral communicative performance of fifth semester college students of French. All practice was conducted using HyperCard stacks and interactive videodisc with the computer providing for the recording of the subjects' oral samples. A maximum of five recording tries per practice program was allowed. The study also investigated the effects of an additional independent variable, task level.

A 2 x 2 factorial design was chosen for the study because of its potential for examining not only the main effects of the two independent variables, subtitling and task level, but also any interaction that might occur. The subtitles availability variable consisted of two levels, "subtitles" or "no-subtitles." Subjects under the "subtitles" condition had video segments subtitled in French on their TV monitor during the task practice. Subjects in
the "no-subtitles" treatment had the same video segments at their disposal, but without subtitles.

The level of the oral transactional tasks, description and narration, was based on the level of processing of the video segments which was required to successfully accomplish those tasks. There were two conditions for this variable, "lower-level" and "higher-level." Descriptive and a narrative lower-level tasks respectively asked subjects to provide, based on what they had seen in the video segments, at least one spatial detail and three specific elements describing the people or narrating the story. Descriptive and narrative higher-level tasks required various spatial details and at least five descriptive or narrative elements of the same people or story. All task directions were provided in English to guarantee that they were understood by all subjects.

Combination of the above variables' levels resulted in the following treatments or cells: T1 (subtitles/lower-level task), T2 (no-subtitles/lower-level task), T3 (subtitles/higher-level task), and T4 (no-subtitles/higher-level task).

Several simultaneous measures were stored by the computer at the various stages of the experimental tasks (viewing video segment, answering "A game"/"A quiz," drafting and recording oral sample). These measures, time on task, oral sample length, "A game"/"A quiz" scores, and branching
and recording frequencies, were designed as descriptive measures, and were not considered part of the experimental design per se.

There were two experimental sessions with two similar programs, Program 1 and Program 2 (thereafter, P1 and P2). During the first session, which was aimed at familiarizing all groups with the workstations and the requirements of the tasks, subjects worked on the unit they chose from P1. During the second session, subjects worked on a unit from P2 similar to the one they had selected during the first session. They also filled out one of the two forms of a Likert questionnaire about their attitudes toward the multimedia courseware package, "Practicing Spoken French" (PSF), used for their speaking practice. The questionnaire forms were primarily designed to provide the subjects' input about the strengths and weaknesses of PSF, and the value of the computerized oral practice experience.

The dependent measure, oral communicative performance, was computed as follows. Each of the oral samples recorded by the subjects during the two practice sessions was given a composite score following the application of the instrument devised by the researcher. This composite score was obtained by weighing by three the score obtained on one of the instrument's six-point subscales, Effectiveness, and adding the scores obtained on the remaining scales, Accuracy, Organization, and Fluency.
Analyses of variance appropriate to one-way and two-way factorial designs were computed using the *SuperANOVA* (1991) software package on a Macintosh Ilsi. The package was also used to run various Tukey-Kramer post-hoc tests to determine the nature of the single interactions that emerged from the ANOVA analyses. Another statistical package, *SPSS* (1986), was also used to calculate the correlation between subjects' oral communicative performance and attitude toward the oral practice experience as well as to determine the inter-rater reliability of the oral performance rating instrument.

**Results**

*Experimental Data*

**Subtitles, Task Level, and Oral Performance**

In order to "double-check" the soundness of the potential effects of the independent variables, the statistical measures used in this study were applied to the experimental and descriptive data gathered from programs P1 and P2, used in the two experimental sessions --even if the first session was only meant to serve as practice for the second session.

The paragraphs that follow describe the data and the results related to the three research questions of this study.
F-Ratios. The first research question asked the following: Does subtitled video vs. unsubtitled video during transactional task practice have any effect on oral communicative performance of fifth semester college students of French? The answer to this primary question is that subtitled video had such an effect: subjects who watched the subtitled video segments during the practice experience scored significantly higher on the oral performance measure than subjects who watched the unsubtitled video segments. The analyses of variance (subtitling x task level) performed on the overall Oral Performance scores of both programs yielded a significant F-ratio for this variable in both P1 (F(1, 40) = 74.60, p< .001), and P2 (F(1, 40) = 68.41, p< .001) (see Tables 5.1 and 5.2).

Also significant at the p <.001 level, were the F-ratios resulting from the analyses of variance (subtitling x task level) performed on the four oral performance subscores of 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)) (see Tables 5.3 to 5.10).

Further illustrating these significant effects, the bar charts in Appendix M (see Figures M.1 to M.5) show the high performance of all cells at the overall and component levels, and, particularly, the
### Table 5.1 Summary ANOVA on Overall Oral Performance Scores (P1) for the Subtitling and Task Level Independent Variables.

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Sum of Squares</th>
<th>Mean Squares</th>
<th>F-Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subtitling</td>
<td>1</td>
<td>297.502</td>
<td>297.502</td>
<td>74.602*</td>
<td>.0001</td>
</tr>
<tr>
<td>Task Level</td>
<td>1</td>
<td>14.316</td>
<td>14.316</td>
<td>3.590</td>
<td>.0654</td>
</tr>
<tr>
<td>Sub. * T. L.</td>
<td>1</td>
<td>.041</td>
<td>.041</td>
<td>.010</td>
<td>.9195</td>
</tr>
<tr>
<td>Residual</td>
<td>40</td>
<td>159.514</td>
<td>3.988</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Dependent: Overall Oral Performance Scores (P1) \( F_{\alpha} \) for \( \alpha = .05 \) (df 1, 40) is 4.08

### Table 5.2 Summary ANOVA on Overall Oral Performance Scores (P2) for the Subtitling and Task Level Independent Variables.

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Sum of Squares</th>
<th>Mean Squares</th>
<th>F-Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subtitling</td>
<td>1</td>
<td>396.651</td>
<td>396.651</td>
<td>68.410*</td>
<td>.0001</td>
</tr>
<tr>
<td>Task Level</td>
<td>1</td>
<td>7.511</td>
<td>7.511</td>
<td>1.295</td>
<td>.2618</td>
</tr>
<tr>
<td>Sub. * T. L.</td>
<td>1</td>
<td>7.024</td>
<td>7.024</td>
<td>1.211</td>
<td>.2776</td>
</tr>
<tr>
<td>Residual</td>
<td>40</td>
<td>231.926</td>
<td>5.798</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Dependent: Overall Oral Performance Scores (P2)
<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Sum of Squares</th>
<th>Mean Squares</th>
<th>F-Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subtitling</td>
<td>1</td>
<td>35.852</td>
<td>35.852</td>
<td>41.284*</td>
<td>.0001</td>
</tr>
<tr>
<td>Task Level</td>
<td>1</td>
<td>2.640</td>
<td>2.640</td>
<td>3.040</td>
<td>.0889</td>
</tr>
<tr>
<td>Sub. * T. L.</td>
<td>1</td>
<td>.001</td>
<td>.001</td>
<td>.001</td>
<td>.9744</td>
</tr>
<tr>
<td>Residual</td>
<td>40</td>
<td>34.736</td>
<td>.868</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Dependent: Effectiveness Scores (P1)

---

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Sum of Squares</th>
<th>Mean Squares</th>
<th>F-Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subtitling</td>
<td>1</td>
<td>76.225</td>
<td>76.225</td>
<td>40.290*</td>
<td>.0001</td>
</tr>
<tr>
<td>Task Level</td>
<td>1</td>
<td>.042</td>
<td>.042</td>
<td>.022</td>
<td>.8819</td>
</tr>
<tr>
<td>Sub. * T. L.</td>
<td>1</td>
<td>.806</td>
<td>.806</td>
<td>.426</td>
<td>.5177</td>
</tr>
<tr>
<td>Residual</td>
<td>40</td>
<td>75.676</td>
<td>1.892</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Dependent: Effectiveness Scores (P2)
Table 5.5  Summary ANOVA on Accuracy Scores (P1) for the Subtitling and Task Level Independent Variables.

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Sum of Squares</th>
<th>Mean Squares</th>
<th>F-Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subtitling</td>
<td>1</td>
<td>23.128</td>
<td>23.128</td>
<td>43.123*</td>
<td>.0001</td>
</tr>
<tr>
<td>Task Level</td>
<td>1</td>
<td>1.319</td>
<td>1.319</td>
<td>2.459</td>
<td>.1247</td>
</tr>
<tr>
<td>Residual</td>
<td>40</td>
<td>21.448</td>
<td>.536</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Dependent: Accuracy Scores (P1)

Table 5.6  Summary ANOVA on Accuracy Scores (P2) for the Subtitling and Task Level Independent Variables.

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Sum of Squares</th>
<th>Mean Squares</th>
<th>F-Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subtitling</td>
<td>1</td>
<td>14.240</td>
<td>14.240</td>
<td>25.882*</td>
<td>.0001</td>
</tr>
<tr>
<td>Task Level</td>
<td>1</td>
<td>.193</td>
<td>.193</td>
<td>.350</td>
<td>.5575</td>
</tr>
<tr>
<td>Sub. * T. L.</td>
<td>1</td>
<td>.547</td>
<td>.547</td>
<td>.994</td>
<td>.3247</td>
</tr>
<tr>
<td>Residual</td>
<td>40</td>
<td>22.008</td>
<td>.550</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Dependent: Accuracy Scores (P2)
Table 5.7  Summary ANOVA on Organization Scores (P1) for the Subtitling and Task Level Independent Variables.

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Sum of Squares</th>
<th>Mean Squares</th>
<th>F-Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subtitling</td>
<td>1</td>
<td>10.756</td>
<td>10.756</td>
<td>23.245*</td>
<td>.0001</td>
</tr>
<tr>
<td>Task Level</td>
<td>1</td>
<td>.761</td>
<td>.761</td>
<td>1.645</td>
<td>.2070</td>
</tr>
<tr>
<td>Sub. * T. L.</td>
<td>1</td>
<td>.399</td>
<td>.399</td>
<td>.863</td>
<td>.3585</td>
</tr>
<tr>
<td>Residual</td>
<td>40</td>
<td>18.509</td>
<td>.463</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Dependent: Organization Scores (P1)

Table 5.8  Summary ANOVA on Organization Scores (P2) for the Subtitling and Task Level Independent Variables.

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Sum of Squares</th>
<th>Mean Squares</th>
<th>F-Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subtitling</td>
<td>1</td>
<td>15.834</td>
<td>15.834</td>
<td>31.690*</td>
<td>.0001</td>
</tr>
<tr>
<td>Task Level</td>
<td>1</td>
<td>1.813</td>
<td>1.813</td>
<td>3.628</td>
<td>.0640</td>
</tr>
<tr>
<td>Sub. * T. L.</td>
<td>1</td>
<td>.697</td>
<td>.697</td>
<td>1.395</td>
<td>.2445</td>
</tr>
<tr>
<td>Residual</td>
<td>40</td>
<td>19.986</td>
<td>.500</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Dependent: Organization Scores (P2)
Table 5. 9  Summary ANOVA on Fluency Scores (P1) for the Subtitling and Task Level Independent Variables.

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Sum of Squares</th>
<th>Mean Squares</th>
<th>F-Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subtitling</td>
<td>1</td>
<td>12.270</td>
<td>12.270</td>
<td>33.804*</td>
<td>.0001</td>
</tr>
<tr>
<td>Task Level</td>
<td>1</td>
<td>.220</td>
<td>.220</td>
<td>.606</td>
<td>.4410</td>
</tr>
<tr>
<td>Sub. * T. L.</td>
<td>1</td>
<td>.037</td>
<td>.037</td>
<td>.103</td>
<td>.7503</td>
</tr>
<tr>
<td>Residual</td>
<td>40</td>
<td>14.517</td>
<td>.363</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Dependent: Fluency Scores (P1)

Table 5. 10  Summary ANOVA on Fluency Scores (P2) for the Subtitling and Task Level Independent Variables.

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Sum of Squares</th>
<th>Mean Squares</th>
<th>F-Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subtitling</td>
<td>1</td>
<td>10.773</td>
<td>10.773</td>
<td>33.384*</td>
<td>.0001</td>
</tr>
<tr>
<td>Task Level</td>
<td>1</td>
<td>.810</td>
<td>.810</td>
<td>2.511</td>
<td>.1209</td>
</tr>
<tr>
<td>Sub. * T. L.</td>
<td>1</td>
<td>.001</td>
<td>.001</td>
<td>.002</td>
<td>.9615</td>
</tr>
<tr>
<td>Residual</td>
<td>40</td>
<td>12.908</td>
<td>.323</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Dependent: Fluency Scores (P2)
outperformance of the subtitles subjects. The overall Oral Performance mean scores of T1 and T3 were respectively 32.25 and 33.46 in P1, and 32.92 and 34.55 in P2 --the highest possible score being 36. Also, the Effectiveness Score of T3 subjects equalized the highest possible score (18.00) in both programs.

The second research question was this one: Do lower- vs. higher-level oral transactional tasks have any effect, not attributable to subtitles, upon oral communicative performance of fifth semester college students of French? On the basis of the data from this study, this question has to be answered negatively. The foregoing analyses of variance indicated non-significant F-ratios for this variable. However, although the F values were not statistically significant, the general pattern of higher-level task subjects outperforming their lower-level task counterparts remained consistent throughout the two programs. Indeed, in P1, the mean scores of the higher-level task subjects on each of the oral communicative performance criteria (see Appendix N, Tables N. 1 to N. 5) were higher than the mean scores of the lower-level tasks subjects. In the same way, in P2, all mean scores of the higher-level task subjects were greater than those of their lower-level task counterparts (see Appendix O, Tables O. 1 to O. 5), with the exception of the Accuracy mean score of T4 subjects which was slightly lower than the mean score of T2 subjects (3.91 vs. 4.0).
The third research question asked this: Is there any interaction between the two independent variables? On the basis of the analyses of variance, the answer to this question is that there was no significant interaction between the subtitling variable and the task level variable. The lines connecting the cell means in the interaction plots (subtitles x task) below (see Figures 5.1 and 5.2) being nearly parallel (within sampling fluctuation), illustrate of this lack of interaction.

**Post-Hoc Tests.** In order to determine which pairs or combinations of means differed for the subtitles main effect, a Tukey-Kramer post-hoc test was computed for both P1 and P2. This test uses the \( Q \) distributions as the sampling distributions, and it is recommended for pairwise comparisons when group sizes are unequal (Steel and Torrie, 1980).

The results of this test, which are provided in Tables 5.11 and 5.12, indicated that in both programs three pairs of combinations of subtitling (A) and task level (B) had a significant influence upon overall Oral Performance. In both programs the subtitles/lower-level task combination \((A_1,B_1)\) was statistically different from both the no-subtitles/lower-level task condition \((A_2,B_1)\) and the no-subtitles/higher-level task condition \((A_2,B_2)\) at the \( p < .05 \) level. The combination subtitles/higher-level task \((A_1,B_2)\) was also significantly different from the no-subtitles/higher-level task condition \((A_2,B_2)\) at the \( p < .05 \) level.
Figure 5.1 Interaction Plot of Overall Oral Performance Scores (P1) per Subtitling and Task Level Effects.

Figure 5.2 Interaction Plot of Overall Oral Performance Scores (P2) per Subtitling and Task Level Effects.
Table 5.11  All Pairwise Comparisons of Means (Tukey-Kramer’s) on Overall Oral Performance Scores (P1) for the Subtitling and Task Level Independent Variables

<table>
<thead>
<tr>
<th>A X B pairs</th>
<th>Mean Differences</th>
<th>(Q_{\text{statistic}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A_1B_1) X (A_2B_1)</td>
<td>5.15</td>
<td>8.58*</td>
</tr>
<tr>
<td>(A_1B_1) X (A_2B_2)</td>
<td>-1.21</td>
<td>-2.05</td>
</tr>
<tr>
<td>(A_1B_1) X (A_2B_3)</td>
<td>4.07</td>
<td>6.90*</td>
</tr>
<tr>
<td>(A_2B_1) X (A_2B_3)</td>
<td>-6.36</td>
<td>-10.26</td>
</tr>
<tr>
<td>(A_2B_1) X (A_2B_2)</td>
<td>-1.08</td>
<td>-1.74</td>
</tr>
<tr>
<td>(A_1B_2) X (A_2B_2)</td>
<td>5.28</td>
<td>8.80*</td>
</tr>
</tbody>
</table>

The \(Q_{cv}\) for \(\alpha = .05, r=4,\) and \(df_w = 40\) is 2.86

\(A_1 = \) Subtitles \(A_2 = \) No-subtitles \(B_1 = \) Lower-level Task \(B_2 = \) Higher-level Task

Table 5.12  All Pairwise Comparisons of Means (Tukey-Kramer’s) on Overall Oral Performance Scores (P2) for the Subtitling and Task Level Independent Variables

<table>
<thead>
<tr>
<th>A X B pairs</th>
<th>Mean Differences</th>
<th>(Q_{\text{statistic}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A_1B_1) X (A_2B_1)</td>
<td>5.22</td>
<td>7.15*</td>
</tr>
<tr>
<td>(A_1B_1) X (A_2B_2)</td>
<td>-1.64</td>
<td>-2.31</td>
</tr>
<tr>
<td>(A_1B_1) X (A_2B_3)</td>
<td>5.19</td>
<td>7.30*</td>
</tr>
<tr>
<td>(A_2B_1) X (A_2B_3)</td>
<td>-6.86</td>
<td>-9.27</td>
</tr>
<tr>
<td>(A_2B_1) X (A_2B_2)</td>
<td>-0.03</td>
<td>-0.04</td>
</tr>
<tr>
<td>(A_1B_2) X (A_2B_2)</td>
<td>6.82</td>
<td>9.34*</td>
</tr>
</tbody>
</table>

The \(Q_{cv}\) for \(\alpha = .05, r=4,\) and \(df_w = 40\) is 2.86

\(A_1 = \) Subtitles \(A_2 = \) No-subtitles \(B_1 = \) Lower-level Task \(B_2 = \) Higher-level Task

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Subtitles, Task Type, and Oral Performance

The additional question of this study, which arose from the combination of the subtitling and task type variables, asked the following:

Does subtitled video vs. unsubtitled video during descriptive vs. narrative transactional task practice have any effect upon oral communicative performance of fifth semester college students of French?

The answer to this question is that there is such an effect. A 2 (subtitles vs no-subtitles) x 2 (description vs narration) ANOVA performed on overall Oral Performance scores of P1 and P2 produced (see Tables 5.13 and 5.14) a significant F-ratio for the subtitles condition in both P1 (F (1, 40) = 69.411, p< .001), and P2 (F (1, 40) = 57.797, p< .001).

The interaction bar charts in Figures 5.3 and 5.4, illustrate the outperformance in both programs of the subtitles subjects on the narrative and the descriptive tasks. In addition, the mean scores of subjects performance on these two types of tasks (see Appendix P, Tables P.1 and P.2) indicate that in both programs, subtitles subjects did better on the narrative task (33.36 (P1); 34.50 (P2)) than on the descriptive task (32.33 (P1); 32.82 (P2)).

The results of two Tukey-Kramer tests, which are provided in Tables 5.15 and 5.16, indicated that in both programs, the subtitles/description combination (A,D) was statistically different from the no-subtitles/description
Table 5.13 Summary ANOVA on Overall Oral Performance Scores (P1) for the Subtitling and Task Type Independent Variables

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Sum of Squares</th>
<th>Mean Squares</th>
<th>F-Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subtitling</td>
<td>1</td>
<td>291.318</td>
<td>291.318</td>
<td>69.411*</td>
<td>.0001</td>
</tr>
<tr>
<td>Task Type</td>
<td>1</td>
<td>2.879</td>
<td>2.879</td>
<td>.686</td>
<td>.4124</td>
</tr>
<tr>
<td>Sub. * T. T.</td>
<td>1</td>
<td>2.879</td>
<td>2.879</td>
<td>.686</td>
<td>.4124</td>
</tr>
<tr>
<td>Residual</td>
<td>40</td>
<td>167.879</td>
<td>4.197</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Dependent: Overall Oral Performance Scores (P1)

Table 5.14 Summary ANOVA on Overall Oral Performance Scores (P2) for the Subtitling and Task Type Independent Variables

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Sum of Squares</th>
<th>Mean Squares</th>
<th>F-Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subtitling</td>
<td>1</td>
<td>331.183</td>
<td>331.183</td>
<td>57.797*</td>
<td>.0001</td>
</tr>
<tr>
<td>Task Type</td>
<td>1</td>
<td>13.151</td>
<td>13.151</td>
<td>2.295</td>
<td>.1377</td>
</tr>
<tr>
<td>Sub. * T. T.</td>
<td>1</td>
<td>2.697</td>
<td>2.697</td>
<td>.471</td>
<td>.4966</td>
</tr>
<tr>
<td>Residual</td>
<td>40</td>
<td>229.203</td>
<td>5.730</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Dependent: Overall Oral Performance Scores (P2)
Figure 5.3 Interaction Bar Chart of Overall Performance Scores (P1) per Subtitling and Task Type Effects.

Figure 5.4 Interaction Bar Chart of Overall Oral Performance Scores (P2) per Subtitling and Task Type Effects.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
### Table 5.15  All Pairwise Comparisons of Means (Tukey-Kramer's) on Overall Oral Performance Scores (P1) for the Subtitling and Task Type Independent Variables.

<table>
<thead>
<tr>
<th>A X D/N pairs</th>
<th>Mean Differences</th>
<th>$Q_{statistic}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A,D X A,D</td>
<td>4.66</td>
<td>7.90*</td>
</tr>
<tr>
<td>A,D X A'N</td>
<td>-1.03</td>
<td>-1.72</td>
</tr>
<tr>
<td>A,D X A'N</td>
<td>4.66</td>
<td>7.28*</td>
</tr>
<tr>
<td>A,D X A'N</td>
<td>-5.69</td>
<td>-9.48</td>
</tr>
<tr>
<td>A,N X A'N</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>A,N X A'N</td>
<td>5.69</td>
<td>8.75*</td>
</tr>
</tbody>
</table>

The $Q_{cv}$ for $\alpha = .05$, $r=4$, and $df_w=40$ is 2.86

$A_1$ = Subtitles  
$A_2$ = No-subtitles  
$D$ = Description  
$N$ = Narration

### Table 5.16  All Pairwise Comparisons of Means (Tukey-Kramer's) on Overall Oral Performance Scores (P2) for the Subtitling and Task Type Independent Variables.

<table>
<thead>
<tr>
<th>A X D/N pairs</th>
<th>Mean Differences</th>
<th>$Q_{statistic}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A,D X A,D</td>
<td>5.29</td>
<td>7.90*</td>
</tr>
<tr>
<td>A,D X A'N</td>
<td>-1.68</td>
<td>-2.37</td>
</tr>
<tr>
<td>A,D X A'N</td>
<td>4.65</td>
<td>5.41*</td>
</tr>
<tr>
<td>A,D X A'N</td>
<td>-6.97</td>
<td>-10.56</td>
</tr>
<tr>
<td>A,D X A'N</td>
<td>-0.64</td>
<td>-0.78</td>
</tr>
<tr>
<td>A,N X A'N</td>
<td>6.33</td>
<td>7.44*</td>
</tr>
</tbody>
</table>

The $Q_{cv}$ for $\alpha = .05$, $r=4$, and $df_w=40$ is 2.86

$A_1$ = Subtitles  
$A_2$ = No-subtitles  
$D$ = Description  
$N$ = Narration

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
(A₂D) and the no-subtitles/narration (A₂N) conditions, at the p < .05 level. The combination subtitles/narration (A₁N) was also significantly different from the combination no-subtitles/narration (A₂N).

**Descriptive Data**

Following the recommendations of Reichardt and Cook (1979), this study makes use of a combination of quantitative and qualitative research paradigms. In addition to the experimental data which provide the basic research evidence, descriptive data were gathered to "round out" this evidence. The data were also intended at providing insights on five CALL issues relevant to the study (locus of control, time on task, learning environment, word-processing, and attitudes). These data included computer and researcher's observations concerning the quantity and quality of users' interaction with the PSF courseware, and the results of the two forms of the attitudes questionnaire dealing with participation and learning materials in this research.

These data should not be construed as experimental results, nor can they legitimately be utilized to predict results in replications using different experimental subtitled video segments and oral tasks.

**Data From Computer Observations**

In all, ten sets of data were collected by the computer for each of the oral tasks practiced by the subjects using the "Practicing Spoken French"
(PSF) courseware. This data collection were unobtrusive and simultaneous with the speaking practice. The ten sets included:

1. total working time (the time spent on one of the units of the PSF programs)
2. initial video processing time (the amount of time spent watching a video segment before answering the related video comprehension questions --"A game" or "A quiz")
3. "A game"/"A quiz" time (the amount of time spent processing the video comprehension questions, deciding of and entering the possible correct answer, and verifying the answer)
4. drafting speech time (the amount of time spent preparing the oral sample. This amount included the time rewatching the video segments and the time typing/handwriting)
5. "Très Utile" time (the time browsing through the "Très Utile" stack)
6. recording time (the amount of time spent on the "Recording Studio")
7. length of recorded oral sample
8. number of times to enter "Très Utile"
9. recordings tries
10. "A game"/"A quiz" scores
Analysis of the computer data indicates that the following trends occurred in this experiment in relation to the subjects’ working times, navigating paths, and "A game"/"A quiz" scores.

**Working Times.** Figure 5.5 shows the means of the total amount of time that subjects spent working on a unit of a PSF program. In light of this information the following remarks are in order.

. The total working time of all cells in P2 was shorter than in P1 -- and yet the overall Oral Performance scores of all cells except T4 were greater in the second than in the first program. The greatest reduction was for T1 subjects who cut their total working time mean by 05:14 minutes. These reductions were most likely due to the practice effects.

. In both programs, subjects assigned to the subtitles/higher-level condition spent more time working than their no-subtitles/higher-level counterparts. Subjects assigned to the subtitles/lower-level condition spent more time working than their no-subtitles/higher-level counterparts in P1 but less than them in P2.

. The difference between the total working times spent in both programs was higher between the higher-level tasks subjects than between the lower-level ones. Indeed, the difference between T3 and T4 total working times was 07:03 for P1 and 08:51 for P2 whereas the difference for the same concept between T1 and T2 was 00:37 for P1.
Figure 5.5 Interaction Bar Chart of Total and Detailed Working Time per Treatment and Program
and 02:16 for P2. These differences could be attributable to the task level variable: the higher-level tasks requiring larger oral output than the lower-level ones, T3 subjects took more time than T4 to process the additional source of information, subtitles, which they had at their disposal to help with the required output.

Figure 5. 5 also shows the means of the amount of time that the subjects spent on each of the "mandatory" stages of the main path of a program unit ("What to See," "A Game" or "A quiz," "Your Draft," "Recording Studio") as well as in the optional path ("Très Utile") of that program unit. Observation of those means leads to the following comments.

- In both programs, the initial video processing times of the subtitles subjects were shorter than those of the no-subtitles subjects.
- In both programs, subjects in the subtitles/lower-level task condition spent less time on the "A game"/ "A quiz" than their no-subtitles/lower-level task counterparts. The opposite was true for the subtitles/higher-level tasks subjects vs. the no-subtitles/higher-level tasks ones.
- In both programs, subtitles subjects spent more time drafting their oral samples than no-subtitles subjects. The drafting speech time of T1 and T2 decreased in P2 as compared with P1 (16:42, 16:30 vs.
18:23, 17:50) while the drafting time of T3 remained the same (21:09) and that of T4 increased (19:05 vs. 18:45).

In both programs, subtitles/higher-level task subjects spent more time in the "Très Utile" stack (13:20, 16:18) than their no-subtitles/higher-level task counterparts (12:12, 12:31). Subtitles/lower-level task subjects spent more time in "Très Utile" than no-subtitles/lower-level task subjects in P1 (15:5 vs. 15:20) but less in P2 (14:25 vs. 17:40). All cells except T1 stayed longer on "Très Utile" in P2 than in P1.

In both programs, subtitles subjects spent more time recording their oral samples than their no-subtitles counterparts. The recording time for all cells except T1, decreased in P2. Such a decrease was probably due to a practice effect: in P2, subjects were already familiar with the recording instructions and they did their recordings quicker than in P1.

In both programs, the oral samples of the subtitles subjects were longer than the ones of the no-subtitles subjects. According to Figure 5.6, the oral sample mean length ranged from the 75:36 seconds of T3 cell in P1 to the 53:73 seconds of T4 in P2. When comparisons are made on the basis of task level, larger differences emerge between the mean length of higher-level task subjects T3 and T4.
(mean difference in P1: 20:63 seconds; mean difference in P2: 19:27 seconds) than between the mean length of lower-level subjects T1 and T2 (mean difference in P1: 06:62 seconds; mean difference in P2: 11:45 seconds). This pattern underscores those pointed out earlier referring to existence of greater differences between the overall Oral Performance scores and total working times of the higher-level task cells than between the performance and times of the lower-level task cells. The pattern also provides additional evidence as to how the task level variable interacted with the subtitles variable particularly in P2—even if such interaction was not statistically significant.

Navigating Paths. The above working times indicate that subjects invariably followed the various stages (viewing video segment, answering "A game"/"A quiz," drafting and recording oral sample) of the "mandatory" path of a program unit. In addition to that evidence, data about the number of times the subjects visited the "Très Utile" stack or recorded the oral sample provide information about the frequency of the subjects' access to an optional path, or the subjects' execution of a multiple-try command.

As concerns the access to "Très Utile," data show that all subjects made use of this optional path. The average number of times that the path was entered increased from P1 to P2 in T1 (from 2.75 to 2.92) and in T4
Figure 5.6 Interaction Bar Chart of Length Oral Sample per Treatment and Program.
(2.55 to 3.27 times); it remained the same in T2 (2.0), and it decreased in T3 (4.64 to 2.45). These numbers indicate that, overall, the higher-level task subjects entered "Très Utile" more than did the lower-level task subjects, and among the former, the subtitles subjects did so more than the no-subtitles ones.

Regarding the number of recording tries, subjects spent at least half the number of times they were allowed to record their oral samples. The average number of recordings for both programs was as follows: T1 (2.42, 3), T2 (2.8, 2.7), T3 (2.36, 2.45), and T4 (2.18, 1.82). According to these figures, lower-level task subjects recorded more times than higher-level task subjects. This may be attributable to the fact that the former were required to complete shorter oral samples, and therefore had the time to rehearse them more often.

"A game"/"A quiz" Scores. Analysis of data about subjects’ performance on video comprehension questions shows that in both programs the subtitles subjects did significantly better than the no-subtitles ones. Indeed, a 2 (subtitles vs no-subtitles) x 2 (description vs narration) ANOVA (see Tables 5. 17 and 5. 18) performed on the scores obtained in the comprehension questions "A game" or "A quiz" yielded a significant F-ratio in both P1 (F (1, 40) = 7.721, p< .05), and P2 (F (1, 40) = 13.322, p< .001).
Table 5.17  Summary ANOVA on Questions ("A Game"/ "A Quiz") Scores (P1) for the Subtitling and Task Type Independent Variables.

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Sum of Squares</th>
<th>Mean Squares</th>
<th>F-Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subtitling</td>
<td>1</td>
<td>35.468</td>
<td>35.468</td>
<td>7.721*</td>
<td>.0083</td>
</tr>
<tr>
<td>Task Type</td>
<td>1</td>
<td>1.861</td>
<td>1.861</td>
<td>.405</td>
<td>.5281</td>
</tr>
<tr>
<td>Residual</td>
<td>40</td>
<td>183.737</td>
<td>4.593</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Dependent: Questions ("A Game"/ "A Quiz") Scores (P1)

Table 5.18  Summary ANOVA on Questions ("A Game"/ "A Quiz") Scores (P2) for the Subtitling and Task Type Independent Variables.

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Sum of Squares</th>
<th>Mean Squares</th>
<th>F-Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subtitling</td>
<td>1</td>
<td>64.806</td>
<td>64.806</td>
<td>13.322*</td>
<td>.0008</td>
</tr>
<tr>
<td>Task Type</td>
<td>1</td>
<td>9.113</td>
<td>9.113</td>
<td>1.873</td>
<td>.2618</td>
</tr>
<tr>
<td>Sub. * T. T.</td>
<td>1</td>
<td>.092</td>
<td>.092</td>
<td>.019</td>
<td>.8911</td>
</tr>
<tr>
<td>Residual</td>
<td>40</td>
<td>194.582</td>
<td>4.865</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Dependent: Questions ("A Game"/ "A Quiz") Scores (P2)
In both programs, those among the subtitles subjects who did "A quiz" ("Narration" Unit) (see Figures 5. 7 and 5. 8) scored higher than those who did "A game" ("Description" Unit). Scores of no-subtitles subjects on "A quiz" were greater than on "A game" only in the second of the programs.

Scores on either "A game" or "A quiz" of subtitles and no-subtitles subjects decreased from P1 to P2 (see Appendix Q, Tables Q. 1 and Q. 2). This decrease could be explained by the fact that the video segments of P2 were slightly more difficult than those of P1 (see difficulty mean ratings of selected video segments, Appendix K).

The Tukey-Kramer tests computed on the ANOVA results (see Tables 5. 19 and 5. 20) showed that, in P1, there were statistically significant differences, at the p< .05 level, between the subtitles/description (A,D) vs. the no-subtitles/narration (A,N) conditions, as well as between the subtitles/narration (A,N) vs. the no-subtitles/narration (A,N) conditions. In P2, the differences were between the subtitles/description (A,D) vs. the no-subtitles/description (A,D) conditions, and, as in P1, between the subtitles/narration (A,N) vs. the no-subtitles/narration (A,N) conditions.

Data From Researcher Observations

A second type of descriptive data emerged from the researcher's systematic observations during the fourteen days of the experiment. The observations included the interaction of subjects with the PSF courseware,
Figure 5.7 Interaction Bar Chart of Questions ("A Game"/"A Quiz") Scores (P1) per Subtitling and Task Type Effects.

Figure 5.8 Interaction Bar Chart of Questions ("A Game"/"A Quiz") Scores (P2) per Subtitling and Task Type Effects.
### Table 5. 19
All Pairwise Comparisons of Means (Tukey-Kramer's) on Questions ("A Game"/ "A Quiz") Scores (P1) for the Subtitling and Task Type Independent Variables

<table>
<thead>
<tr>
<th>A X D/N pairs</th>
<th>Mean Differences</th>
<th>( Q_{\text{statistic}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( A_1D \times A_2D )</td>
<td>1.34</td>
<td>2.16</td>
</tr>
<tr>
<td>( A_1D \times A_1N )</td>
<td>-0.06</td>
<td>-0.10</td>
</tr>
<tr>
<td>( A_1D \times A_2N )</td>
<td>2.23</td>
<td>3.32*</td>
</tr>
<tr>
<td>( A_2D \times A_1N )</td>
<td>-1.40</td>
<td>-2.22</td>
</tr>
<tr>
<td>( A_2D \times A_2N )</td>
<td>0.89</td>
<td>1.33</td>
</tr>
<tr>
<td>( A_1N \times A_2N )</td>
<td>2.29</td>
<td>3.37*</td>
</tr>
</tbody>
</table>

The \( Q_c \) for \( \alpha = .05, r=4, \) and \( df_w = 40 \) is 2.86

\( A_1 = \) Subtitles  \( D = \) Description
\( A_2 = \) No-subtitles  \( N = \) Narration

---

### Table 5. 20
All Pairwise Comparisons of Means (Tukey-Kramer's) on Questions ("A Game"/ "A Quiz") Scores (P2) for the Subtitling and Task Type Independent Variables

<table>
<thead>
<tr>
<th>A X D/N pairs</th>
<th>Mean Differences</th>
<th>( Q_{\text{statistic}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( A_1D \times A_2D )</td>
<td>2.47</td>
<td>3.98*</td>
</tr>
<tr>
<td>( A_1D \times A_1N )</td>
<td>-1.06</td>
<td>-1.63</td>
</tr>
<tr>
<td>( A_1D \times A_2N )</td>
<td>1.60</td>
<td>2.03</td>
</tr>
<tr>
<td>( A_2D \times A_1N )</td>
<td>-3.53</td>
<td>-5.89</td>
</tr>
<tr>
<td>( A_2D \times A_2N )</td>
<td>-0.87</td>
<td>-1.16</td>
</tr>
<tr>
<td>( A_1N \times A_2N )</td>
<td>2.66</td>
<td>3.41*</td>
</tr>
</tbody>
</table>

The \( Q_c \) for \( \alpha = .05, r=4, \) and \( df_w = 40 \) is 2.86

\( A_1 = \) Subtitles  \( D = \) Description
\( A_2 = \) No-subtitles  \( N = \) Narration

---

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
their reactions to the experimental site and comments on PSF, and their oral sample drafts and transcripts.

Qualitatively, the subjects' interaction with the "Practicing Spoken French" courseware was greater than expected, given the subjects' unfamiliarity with HyperCard. In the first session, all but one of the subjects required some kind of assistance to continue their work on the selected practice unit, either because they had a "real" problem or simply because they were afraid to touch the "wrong" key. In the second session, however, most of the subjects went through their units without any help.

Quantitatively, the subjects's interaction with PSF was high and increased from the first to the second experimental session. All subjects accomplished the four stages of the experimental tasks (viewing, checking comprehension input, drafting and recording one's own oral sample) although some of them disregarded the instructions of the "PSF User" stack's cards and help windows. During the first session, for instance, two subjects overlooked the direction in the card which asked them to click on the Scores button to see a how well they had performed in the "A game" questions. Similarly, during the second practice session, nine subjects (three from T1, one from T2, three from T3, and two from T4) switched from the unit they had worked on during the previous session to the alternative
unit because they had skipped the directions found on the help window of the "Home" card.

Reactions to the experimental site were quite positive. Most of the subjects seemed to be able to work at ease without being distracted by the presence of other people in the computer laboratory. Some of them, however, complained about the furniture (the tables were too low) or the talk of students from another course who were working on a team project.

Reactions to the program were expressed not only through answers to the attitude questionnaires but also through spontaneous comments during the experimental sessions. Sample positive comments follow:

"That's neat!"

"That's very fun!"

"That's cool!"

"I would like to come again and practice with the other units."

"I would like to buy this program."

"I have learnt very useful things in "Très Utile."

"I find the program very clear and interesting. The grammar programs at the language lab are boring."

"I have a feeling of achievement with this program."

Negative comments were also heard such as the one from the student who disliked the program because she didn't want "to hear her own
Other students, however, who at the beginning of the first session appeared to have a negative attitude toward computers, not only changed that attitude but also improved the quality of their productions from the first to the second session.

Finally, observation of the subjects' oral sample drafts and transcripts led to interesting findings concerning the effects of word processing and the use of lexical items seen in the "Très Utile" stack or the subtitled video segments. First, 54% of the subjects used the computer to type their oral sample drafts during the first session, and 58% during the second one. These percentages were surprisingly high given the fact that the majority of the students were neither experienced Macintosh users nor skilled word processors. But, the most interesting phenomenon was that those subjects who typed their drafts on the computer organized the content better and made fewer spelling errors than subjects who did them by hand. Word processing seemed to increase subjects' concern for accuracy even if it had been told that the oral sample drafts were not going to be evaluated.

Second, oral sample drafts and transcripts also showed how much subjects had interacted with the "Très Utile" stack and the video subtitles, and how had they reused the terms seen there. It was observed that, in all the instructional units, at least five words of the "Difficult Words" section of "Très Utile" had been reemployed by the majority of the subjects. Reemployment
of the connectors seen in the "How to Narrate" section of that stack was also very high, particularly among the subtitles/higher-level task subjects. It was equally observed that reemployment of difficult words heard in the video segments, such as numbers and proper names, was higher among the subtitles than the no-subtitles subjects.

**Data From Attitude Questionnaires**

A third type of data gathered during the course of this study included the results of two Likert questionnaires. At the end of the second practice session, two questionnaires, form A for the subtitles treatments, and form B for the no-subtitles ones, were administered to evaluate subjects' attitudes toward their speaking practice experience with multimedia courseware (see Appendices D and E).

Since subjects were told before the experiment that they would be asked to evaluate the practice materials, they took very seriously their role of "evaluators." Not only did they answer carefully the questionnaires but also showed their interest in providing suggestions about improving PSF's "Drafting Card" by adding a spelling checker, or about including alternative video segments.

Overall, the subjects demonstrated a highly positive attitude toward the speaking practice with multimedia courseware, subtitles subjects being significantly more in favor of this kind of practice than no-subtitles subjects.
Indeed, the one-way (subtitles vs. no-subtitles) ANOVA performed on attitude ratings revealed a significant subtitles effect (F-ratio (1, 42) = 4.249, p< .05) (see Table 5. 21).

Table 5. 21  Summary ANOVA on Attitude Ratings for the Subtitling Variable

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Sum of Squares</th>
<th>Mean Squares</th>
<th>F-Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subtitling</td>
<td>1</td>
<td>42.646</td>
<td>42.646</td>
<td>4.249*</td>
<td>.0455</td>
</tr>
<tr>
<td>Residual</td>
<td>42</td>
<td>421.536</td>
<td>10.037</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Dependent: Attitude Ratings  The F_0.05 for α = .05 (df 1, 42) is 4.07

Tables 5. 22 and 5. 23 provide the mean scores for the responses of subtitles and no-subtitles subjects to the statements of the attitude questionnaires. The overall mean of the T1 and T3 subjects on the ten item questionnaire form A was 4.63 out of a range of 1.0 - 5.0. The overall mean of the T2 and T4 subjects on the ten item questionnaire form B was 4.44.

Among the statements that addressed the multimedia courseware practice issue, statement 1, which read "Overall, I found the French oral
Table 5.22  Mean Statement Scores of Subtitles Cells (T1, T3) on Attitude Questionnaire Form A

<table>
<thead>
<tr>
<th>Treatment</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>(O.Mean)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subtitles /Lower-Level Task</td>
<td>4.75</td>
<td>4.5</td>
<td>4.5</td>
<td>4.67</td>
<td>4.75</td>
<td>4.83</td>
<td>4.75</td>
<td>4.92</td>
<td>4.42</td>
<td>4.75</td>
<td>4.68</td>
</tr>
<tr>
<td>Subtitles /Higher-Level Task</td>
<td>4.82</td>
<td>4.45</td>
<td>4.64</td>
<td>4.64</td>
<td>4.36</td>
<td>4.45</td>
<td>4.45</td>
<td>4.91</td>
<td>4.0</td>
<td>5.0</td>
<td>4.57</td>
</tr>
<tr>
<td>Overall Means</td>
<td>4.79</td>
<td>4.48</td>
<td>4.57</td>
<td>4.66</td>
<td>4.56</td>
<td>4.64</td>
<td>4.60</td>
<td>4.92</td>
<td>4.21</td>
<td>4.88</td>
<td>4.63</td>
</tr>
</tbody>
</table>
Table 5.23 Mean Statement Scores of No-Subtitles Cells (T2, T4) on Attitude Questionnaire Form B

<table>
<thead>
<tr>
<th>Treatment</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10 (0. Mean)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No-Subtitles/Lower-Level Task</td>
<td>4.3</td>
<td>4.1</td>
<td>4.2</td>
<td>4.4</td>
<td>4.5</td>
<td>4.3</td>
<td>4.7</td>
<td>4.6</td>
<td>4.4</td>
<td>4.8</td>
</tr>
<tr>
<td>No-Subtitles/Higher-Level Task</td>
<td>4.63</td>
<td>4.45</td>
<td>4.45</td>
<td>4.64</td>
<td>4.18</td>
<td>4.54</td>
<td>4.54</td>
<td>4.54</td>
<td>4.18</td>
<td>4.36</td>
</tr>
<tr>
<td>Overall Means</td>
<td>4.46</td>
<td>4.28</td>
<td>4.33</td>
<td>4.52</td>
<td>4.34</td>
<td>4.42</td>
<td>4.62</td>
<td>4.57</td>
<td>4.29</td>
<td>4.58</td>
</tr>
</tbody>
</table>
practice sessions with multimedia courseware to be a worthwhile learning experience," obtained the highest rating in both forms of the questionnaire -- form A = 4.79, and form B = 4.46.

Among the statements concerned with the PSF design, the highest rating in questionnaire form A (4.92) was for statement 8 which read "The subtitles increased my comprehension of what was said on the video segments, and therefore, my ability to express the content of these segments." The highest rating for the same concept in questionnaire form B (4.62) was for statement 7 which read "The information on the 'Très utile' stack provided me the vocabulary and grammar I needed to prepare the descriptive or the narrative tasks." The lowest ratings in both forms of the questionnaire, A = 4.21 and B = 4.29, were for the two versions of statement 9 which respectively asked about the subtitles fonts and video quality.

A comparison of attitude ratings on the basis of percentage of distribution of the questionnaire values ("strongly disagree" ... "strongly agree") shows that more varied and higher rating values were given by subtitles subjects than by no-subtitles subjects. According to the bar charts that follow (see Figures 5. 9 and 5. 10), the rating values given by subjects who took questionnaire form A ranged from "disagree" to "strongly agree" whereas the values accorded by subjects who filled out questionnaire form
Figure 5.9 Bar Chart of Overall Reactions of Subtitles Cells (T1, T3) to Attitude Questionnaire Form A
Figure 5.10 Bar Chart of Overall Reactions of No-Subtitles Cells (T2, T4) to Attitude Questionnaire Form B
B ranged from "no opinion" to "strongly agree." Furthermore, every statement in form A obtained a higher percentage of "strongly agree" ratings than found in form B. Statement 8 of questionnaire A, which asked about the effectiveness of the subtitles, was rated with "strongly agree" by 91.30% of the subtitles subjects whereas statement 8 of questionnaire form B, which asked about the computer controlling features, received the "strongly agree" rating of 61.90% of the no-subtitles subjects. In the same way, statement 10 of both forms, which read "The time and the number of tries allowed to record the description or the narration were sufficient," was rated with "strongly agree" by 86.95% of the subtitles subjects, whereas only 57.14% of the no-subtitles subjects accorded this rating to that statement.

Although subjects in all cells showed high positive attitudes toward using multimedia courseware, and more especially the subtitles subjects, the correlation between oral performance and attitude was low to little. Pearson product-moment correlations of attitude ratings with Oral Performance scores of subjects in all cells for P2 were as follows: $T_1 = 0.0573; T_2 = 0.0655; T_3 = 0.4093; T_4 = 0.2050$. According to these figures, the higher correlation, even if it was not significant, corresponded to the subtitles/higher-level task cell.
Summary

The results from the experimental and descriptive data commented in this chapter can be summarized as follows:

- Subtitles had a significant effect on oral communicative performance of fifth semester college students of French. In the two experimental sessions, the overall Oral Performance scores and subscores (Effectiveness, Accuracy, Organization, and Fluency) of subtitles subjects were significantly higher than those of no-subtitles subjects.

- Neither the effects of task level nor the interaction of this variable with the subtitling 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.

- The combination subtitles/higher-level task turned out to be the most beneficial within and across time. In both programs, T3 subjects obtained the highest scores in the oral performance communicative measure. Moreover, the differences in P1 between overall Oral Performance mean scores, total working time, and oral sample length of T3 and T4 were higher than between T1 and T2, and those differences increased on P2.
Subtitles had a significant effect on overall Oral Performance scores when combined with task type, this effect being higher with the narrative than with the descriptive tasks.

Subtitles/higher-level task subjects spent more time working in both programs than subjects on the other three cells. And the working times of all cells decreased from P1 to P2.

The courseware's "mandatory" navigating path was invariably followed by all subjects. The optional path "Très Utile" was also frequented by all subjects who stayed there longer in P2 than in P1. The path was entered more by the higher- than by the lower-level tasks subjects, and among the former, more by the subtitles than by the no-subtitles ones. The recording of the oral sample was, however, performed more often by the lower- than by the higher-level task subjects.

Subtitles also affected subjects performance on the video comprehension questions "A game" and "A quiz." Subtitles subjects scored significantly higher than no-subtitles subjects on those questions, and they did better on "A quiz" than on "A game."

Researcher's observations indicated that word processing affected the organization and formal accuracy of the oral sample drafts. The observations also shed additional light over the quantity and the quality of the interaction of subjects with the experimental courseware.
"Practicing Spoken French," and the reactions of subjects to the
experimental site.

- Attitudes of all subjects toward the computerized speaking practice
  experience (experience itself and multimedia package) were highly positive.
  Particularly, attitudes of subtitles subjects were significantly superior to
  attitude of no-subtitles subjects. The correlation between attitudes and oral
  performance of subjects in all cells was, however, insignificant.
CHAPTER 6
SUMMARY, CONCLUSIONS, IMPLICATIONS, AND RECOMMENDATIONS

Summary of the Study

Goals
The main objective of this study was to investigate the effects of subtitling during transactional task practice with multimedia courseware on oral communicative performance of fifth semester college students of French. Subtitling refers to whether or not subjects can see a video segment with fully duplicating intralingual subtitles. In addition to subtitling, the independent variable of primary interest, another independent variable, the level of the practice task was also investigated.

Salomon's (1979) theory about the interaction of media, cognition, and learning as well as Salomon's (1981) Amount of Invested Mental Effort (AIME) model provided the theoretical base for the choice of the independent and dependent variables. Following Salomon's advice, the
study investigated how the presence/absence of subtitles, and the higher/lower demands of a learning task work together to affect some aspect of learning, which in this case was the dependent variable, oral communicative performance.

To verify any main effect or interaction, the two independent variables were combined in a 2 x 2 factorial research design.

Literature

Within the framework of Salomon's interactional theory, it was hypothesized that the use of a coding element such as subtitles would shortcut skills such as listening and reading, and thus spare processing capacity in favor of the semantic encoding and delivery of information. Although this hypothesis hasn't been tested by CALL basic studies, it is largely supported by theoretical literature on cognitive psychology (serial-parallel processing, bottom-up-top-down processing, and between channel redundancy).

In the same way, within the confines of Salomon's AIME model, it was thought that the manipulation of the quantity and depth of the information input required to accomplish the learning task would affect learner's perception of the task and, subsequently, learner's performance. Research on instructional media supporting this hypothesis was presented.

Also, findings of research on the sciences about the effects of task structure on performance were discussed and taken into account in
designing the multimedia package, "Practicing Spoken French" (PSF), used for the experimental speaking practice. Insights about how learning materials can be organized to better reflect and shape learner's internal knowledge were applied to PSF in order to avoid "disorientation," a potential drawback to learning when using multimedia tools.

Finally, literature on language testing relevant to the dependent variable, oral communicative performance, as well as literature on CAI and CALL issues indirectly addressed by the study, was also examined.

**Procedures**

Forty-four subjects were assigned to four treatment conditions resulting from the combinations of the subtitling and task level variables. Using the PSF package designed by the researcher, subjects worked during two sessions on two optional units. In each unit, they were requested: 1) to watch a video segment, with or without subtitles, at least two times; 2) to answer various questions about information in that video segment; 3) to write a draft of a description or a narration following the requirements of the task corresponding to the treatment to which they had been ascribed; 4) to record an up to 3-minute oral sample, based on what they had drafted. At the end of the second session, subjects also filled out one of two forms of a Likert questionnaire aimed at evaluating their attitudes toward the speaking practice and the totality of the PSF package. Transcripts of 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 arrived at by adding the individual scores obtained from the four six-point scales (Effectiveness, Accuracy, Organization, and Fluency) of the evaluating instrument developed by the researcher.

Experimental and descriptive data were gathered in four ways: 1) from the transcripts of subjects' recorded oral samples, 2) from the measures unobtrusively recorded by the computer, 3) from the researcher's observation of subjects' oral sample drafts and transcripts as well as subjects' reactions to the experimental sessions, 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 further verify the soundness of the experimental results.

Appropriate ANOVA, post-hoc tests, and Pearson-r were utilized to analyze the experimental data. In addition, descriptive data were analyzed and used by the researcher to contextualize the experimental results, to provide useful directions for future research, and to add new insights on certain CALL issues.
Further Limitations

In addition to the limitations discussed in Chapter 1, other limitations emerged during the implementation of the research. These limitations were of two kinds: external, those related to the hardware or the software used for its implementation, and internal, those related to its design.

Basically, two external limitations were found. First, given that the computers used for the experiment had only 5mb of Random Access Memory (RAM), the allotted time for the recording of the oral samples had to be restricted to three minutes. This restriction had a bearing upon the quality of the samples: because of time pressure, some of the subjects spoke too hastily or skipped some of the ideas they had written down in their drafts.

Second, the fonts of the subtitled video segments, 95 Helvetica and 96 Helvetica Oblique, which were used to distinguish between the words of the narrators and interviewers, and those of the people interviewed, were too similar to make that distinction noticeable. Limitations of the equipment used to generate the subtitles prevented the researcher from selecting more appropriate fonts.

The main internal limitation of the study was probably its sample size. Although sixty subjects were initially targeted to take part in the experiment, only 44 students (12 in T1; 10 in T2; 11 in T3, and 11 in T4) actually
participated. Since sample size is bound to affect the statistical power of a test, this limitation may have influenced the lack of significance of the effects of task level, the second of the independent variables of this study.

**Summary of the Findings**

**Main Findings**

Analysis of the experimental data revealed statistically significant effects for subtitling, the independent variable of primary interest (see Tables 6.1 and 6.2). Indeed, compared to the no-subtitles condition, the subtitles condition resulted in significantly higher overall Oral Performance scores in both P1 \( (F (1, 40) = 74.60, p< .001) \), and P2 \( (F (1, 40) = 68.41, p< .001) \).

The subtitles condition also resulted in four significantly higher subscores in the two experimental 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 demonstrated that in both programs the subtitles/lower-level task combination was statistically different 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
Table 6.1 Significant F-Ratios for the Subtitling Variable Yielded by the ANOVA Analyses

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>Dependent Variable</th>
<th>F-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Program 1</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subtitling/Task Level</td>
<td>Overall Oral Performance</td>
<td>74.602* (p = .0001)</td>
</tr>
<tr>
<td>Subtitling/Task Level</td>
<td>Effectiveness</td>
<td>41.284* (p = .0001)</td>
</tr>
<tr>
<td>Subtitling/Task Level</td>
<td>Accuracy</td>
<td>43.132* (p = .0001)</td>
</tr>
<tr>
<td>Subtitling/Task Level</td>
<td>Organization</td>
<td>23.245* (p = .0001)</td>
</tr>
<tr>
<td>Subtitling/Task Level</td>
<td>Fluency</td>
<td>33.808* (p = .0001)</td>
</tr>
<tr>
<td>Subtitling/Task Type</td>
<td>Overall Oral Performance</td>
<td>69.411* (p = .0001)</td>
</tr>
<tr>
<td>Subtitling/Task Type</td>
<td>&quot;A Game&quot;/&quot;A Quiz&quot; Scores</td>
<td>7.721* (p = .0083)</td>
</tr>
<tr>
<td><strong>Program 2</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subtitling/Task Level</td>
<td>Overall Oral Performance</td>
<td>68.410* (p = .0001)</td>
</tr>
<tr>
<td>Subtitling/Task Level</td>
<td>Effectiveness</td>
<td>40.290* (p = .0001)</td>
</tr>
<tr>
<td>Subtitling/Task Level</td>
<td>Accuracy</td>
<td>25.882* (p = .0001)</td>
</tr>
<tr>
<td>Subtitling/Task Level</td>
<td>Organization</td>
<td>31.690* (p = .0001)</td>
</tr>
<tr>
<td>Subtitling/Task Level</td>
<td>Fluency</td>
<td>33.384* (p = .0001)</td>
</tr>
<tr>
<td>Subtitling/Task Type</td>
<td>Overall Oral Performance</td>
<td>57.797* (p = .0001)</td>
</tr>
<tr>
<td>Subtitling/Task Type</td>
<td>&quot;A Game&quot;/&quot;A Quiz&quot; Scores</td>
<td>13.332* (p = .0083)</td>
</tr>
<tr>
<td>Subtitling</td>
<td>Attitudes Ratings</td>
<td>4.249* (p = .0455)</td>
</tr>
</tbody>
</table>
Table 6.2 Significant Q-Statistics Yielded by the Tukey-Kramer Tests Performed on the ANOVA Results

<table>
<thead>
<tr>
<th>Variables</th>
<th>Pairwise Comparisons</th>
<th>Q_{statistic}</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Program 1</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Independent:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subtitling/Task Level</td>
<td>Subtitles/Lower-level task vs. No-subtitles/Lower-level task</td>
<td>8.58*</td>
</tr>
<tr>
<td>Dependent:</td>
<td>Subtitles/Lower-level task vs. No-subtitles/Higher-level task</td>
<td>6.90*</td>
</tr>
<tr>
<td>Overall Oral Perfor.</td>
<td>Subtitles/Higher-level task vs. No-subtitles/Higher-level task</td>
<td>8.80*</td>
</tr>
<tr>
<td><strong>Independent:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subtitling/Task Type</td>
<td>Subtitles/Description vs. No-subtitles/Description</td>
<td>7.90*</td>
</tr>
<tr>
<td>Dependent:</td>
<td>Subtitles/Description vs. No-subtitles/Narration</td>
<td>7.28*</td>
</tr>
<tr>
<td>Overall Oral Perfor.</td>
<td>Subtitles/Narration vs. No-subtitles/Narration</td>
<td>8.75*</td>
</tr>
<tr>
<td><strong>Independent:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subtitling/Task Type</td>
<td>Subtitles/Description vs. No-subtitles/Narration</td>
<td>3.32*</td>
</tr>
<tr>
<td>&quot;A Game&quot;/ &quot;A Quiz&quot; Sc.</td>
<td>Subtitles/Narration vs. No-subtitles/Narration</td>
<td>3.37*</td>
</tr>
</tbody>
</table>

| **Program 2**              |                                                           |               |
| **Independent:**           |                                                           |               |
| Subtitling/Task Level      | Subtitles/Lower-level task vs. No-subtitles/Lower-level task | 7.15*         |
| Dependent:                 | Subtitles/Lower-level task vs. No-subtitles/Higher-level task | 7.30*         |
| Overall Oral Perfor.       | Subtitles/Higher-level task vs. No-subtitles/Higher-level task | 9.34*         |
| **Independent:**           |                                                           |               |
| Subtitling/Task Type       | Subtitles/Description vs. No-subtitles/Description        | 7.90*         |
| Dependent:                 | Subtitles/Description vs. No-subtitles/Narration          | 5.41*         |
| Overall Oral Perfor.       | Subtitles/Narration vs. No-subtitles/Narration            | 7.44*         |
| **Independent:**           |                                                           |               |
| Subtitling/Task Type       | Subtitles/Description vs. No-subtitles/Description        | 3.98*         |
| "A Game"/ "A Quiz" Sc.     | Subtitles/Narration vs. No-subtitles/Narration            | 3.41*         |

* All the significant differences were in favor of the treatments including the subtitles condition.
also significantly different from the no-subtitles/higher-level task condition at the $p < .05$ level.

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. It was equally found that in both programs subjects in the subtitles/higher-level task cell 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 ($F(1, 40) = 69.411, p < .001$), and P2 ($F(1, 40) = 57.797, p < .001$). And, 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/description condition was statistically different from the no-subtitles/description and the no-subtitles/narration conditions, at the $p < .05$ level. The combination subtitles/narration was also significantly different from the combination no-subtitles/narration.
Additional Findings

An analysis of the data gathered from the measures recorded by the computer, the observations of the researcher, and the attitude questionnaires provided additional light on the experimental results of this study.

Computer Observations

Computer data about subjects' total working time 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 and yet the overall Oral Performance scores of all cells, except T4, increased from the first to the second program.

Computer data about subjects' detailed working times indicated the following. 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 ones; and among the former, the subtitles subjects more so than the no-subtitles ones. Data also revealed that the lower-level task subjects recorded their samples more often than did the higher-level task subjects.

Subtitles had a significant effect on the video comprehension questions, "A game" or "A quiz," which were respectively asked in the "Description" and "Narration" units. Statistical analyses of the subjects scores on these questions indicated that, compared with the no-subtitles subjects, subtitles subjects performed significantly better in the two types of questions, in both P1 (F (1, 40) = 7.721, p < .05), and P2 (F (1, 40) = 13.322, p < .001). Also in both programs, scores of those among the subtitles subjects who did the "A quiz" questions were higher than those who did the "A game" questions.

Tukey-Kramer tests on the results of the comprehension questions showed that, in P1, there were statistically significant differences, at the p < .05 level, between the means of the subtitles/description vs. the no-
subtitles/narration subjects, as well as between those of the subtitles/narration vs. the no-subtitles/narration subjects. In P2, the differences were between the mean questions of the subtitles/description vs. the no-subtitles/description subjects, and, as in P1, between those of the subtitles/narration vs. the no-subtitles/narration subjects.

**Researcher Observations**

The four kinds of observations recorded by the researcher during the two experimental sessions 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 required 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 seemed to like the computer laboratory where the experiment took place better than 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. The spontaneous comments that the subjects made of the PSF courseware during the experimental sessions reinforced the opinions that they expressed on this matter in the attitude questionnaires.
Third, observation of the subjects' oral sample drafts and transcripts revealed that: 1) oral sample drafts that had been typed on the computer (54% in P1, and 58% in P2) were generally better organized and had fewer spelling errors than drafts which had been written by hand; 2) reemployment 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) reemployment of difficult words such as numbers or proper names heard in the video segments was, as it 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, particularly, toward the "Practicing Spoken French" package. This attitude turned out to be significantly higher for the subtitles subjects than for the no-subtitles ones (F-ratio (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 and CALL Issues**

The experimental and especially the descriptive data collected during this study indirectly contributed to research knowledge on the following
CALL issues: 1) the effects of locus of control on performance; 2) the informational value of time on task; 3) the effects of learning environment on attitudes; 4) the effects of word processing on written production quality; and 5) the relationship between attitude toward and performance with multimedia courseware.

Locus of Control

The design of the courseware package, "Practicing Spoken French," used in this study utilized a combined learner/program control formula which, 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, and the textual information resources. Learners could also elect 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 those, combined learner-control and learner-control with advisement, featured by the materials created for previous CALL (Robinson, 1991) and CAI (Johansen & Tennyson, 1983; Arnone & Grabowski, 1992) studies. 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 insuring that their exposure to the units was sufficient to detect any experimental effect.

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 actions which 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 (Orlansky, 1983) and, more importantly, of the learning strategies used by subjects during CALL practice (Pederson, 1985). 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 which had been planned into its design --analyzing, creating structure for input and output, and, practicing.

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, and, 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, particularly, of the practice strategy. Data about the total and detailed working times of both programs indicated that all subjects consistently worked in both the main and the optional paths of the PSF unit for at least the minimal time the researcher had estimated they would spend in those paths.

Learning Environment

Descriptive data from this study also support the findings of CALL research on two important aspects of the learning environment issue, experimental site and learning materials.

As in Robinson et al.'s (1985) study, and as mentioned earlier, subjects in this investigation seemed to prefer the computer laboratory where the experimental sessions took place to the language laboratory. This preference, verbalized in subjects' spontaneous comments, can probably be attributed to three factors: 1) the appeal of the experimental materials which, different from those used in the language laboratory, constituted a "discovery" for the majority of them; 2) the role of "evaluators" which the subjects assumed they were playing; and 3) the pluridisciplinarity of the site which seemed to confer to a humanities subject matter such as French the "prestige" traditionally reserved to scientific disciplines which have had the hegemony of the computer laboratory.
Also, like in Lozano et al.'s (1985) and Blomeyer's (1985) studies the majority of subjects had very high praise for the experimental package. As in those studies, this high level of opinion was partially due to the fact that subjects knew that "Practicing Spoken French" had been developed by a researcher at LSU, for students at LSU.

**Word Processing**

Second/foreign language research has shown the impact of word processing on the quality of written production and on attitudes toward the writing process (Sommers, 1985; Neu & Scarcella, 1991). Results from this study underscore those findings. Indeed, researcher's observations of the subjects' oral sample drafts indicated that subjects who used word processing were more organized and accurate than those who wrote them 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 contributed to the existing empirical knowledge about attitude toward CAI and CALL. As in previous studies (Roblyer et al., 1988), and as mentioned earlier, subjects' attitudes toward the speaking practice with PSF were very positive but correlated poorly with subjects' oral performance. Contrary to Ahmad et al.'s (1986) findings, subjects in this
study with little experience in using computers reacted as favorably toward the computerized speaking practice as did the more experienced subjects.

Conclusions

In light of the above findings, two sets of conclusions are offered, the first regarding the independent variables and the second concerning the experimental courseware.

Independent Variables

Subtitling

The results from this study support and advance the knowledge generated by second/foreign 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 not only to 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 leads to the following main conclusion: when learning from "authentic video," 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 (1979) theoretical propositions, discussed in Chapter 1, about how media symbol systems and coding elements affect cognition and learning. Within the context of those propositions and in light of the experimental results, it can now be concluded that:

1) video with fully duplicating intralingual subtitles may help the foreign/second language learner to 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 interactive technologies such as IVD, may contribute to the internalization of this coding element and the short-cutting of the skill, simultaneous listening and reading, that it facilitates;

3) the simultaneous activation by the subtitles of listening and reading skills 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 scored significantly higher than no-subtitles subjects in the "A quiz" and the "A game" questions as well as in the descriptive and the narrative tasks
provides further indication that subtitles may enhance both listening comprehension and oral production.

Also, the fact that the subtitles subjects who did the narrative tasks and the "A quiz" questions performed better than those who did the descriptive tasks and the "A game" questions allows us to conclude that subtitles may be particularly beneficial with video learning activities that require the subjects to rely more largely on the aural than on the visual information channel.

Next, the fact that subtitles subjects spent about the same or less\(^2\) time than their no-subtitles counterparts watching the video segments at the beginning of the program indicated that they didn't pay too 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, since American audiences rarely watch subtitled foreign movies, leads to the following conclusion: practice\(^3\) in the use of native language subtitles in informal learning settings may improve the attitudes toward and the benefits of the use of target language subtitles in academic foreign/second language settings.

Finally, the fact that the attitudes of subtitles subjects were significantly higher than those of no-subtitles subjects leads to the conclusion that subtitles may increase subjects Perceived Self-Efficacy
(PSE) (Salomon, 1981), 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 scriptural transcript of the aural source.

**Task level**

Although the effects of the secondary independent variable, task level, were insignificant, a pattern of better performance for the higher-level task subjects was found in both programs. The pattern, which underscores results of previous research (Salomon & Leigh, 1984; Krendl & Watkins, 1983) about the effects of perception of task difficulty on mental effort and performance, could have probably been significant if, as mentioned earlier, the sample size had been greater.

Similarly, 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, T3 subjects seemed to retain and 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.

Experimental Courseware

Based on the high performance of all subjects on both programs of the multimedia package "Practicing Spoken French" (PSF), it can be concluded that computerized oral communicative practice is feasible and effective. The effectiveness of PSF may be attributable to the following:

1) PSF provided the subjects with subtitled video segments, a feature which could eliminate the "shallowness" of processing (Singer, 1980; Salomon, 1981), one of the limitations of mediated information;

2) PSF offered to the subjects hierarchically, yet flexibly, structured practice tasks aimed at "modelling" cognitive processes (Eylon & Reif, 1984) and improving oral performance. The PSF components (computer text, pictures, animations, and video and audio segments) were ordered to facilitate the accomplishment of the four stages of a program unit (viewing, checking comprehension input, drafting one own's oral sample, recording one own's oral sample). Subjects had to complete the stages in the specified order but they were free to watch or listen to the video and audio segments or to enter "Très Utile";

3) PSF created a meaningful context within which the subjects could review and practice grammar and vocabulary. The goal-oriented nature
(Chun & Brandl, 1992) as well as the hierarchical structure of the grammatical and lexical information presented in the "Très Utile" stack may have incited subjects, especially the subtitles ones, to take advantage of this optional path;

4) PSF provided the subjects with a tension-free environment which, like in previous studies (Robinson et al., 1985), may have contributed to achievement. As mentioned in Chapter 4, this feature was posited in four ways: by referring to people or locations that the subjects knew in order to involve them personally; by using irony and emotion; by "speaking" to the user with a friendly tone; and by maximizing opportunities for practice;

5) PSF motivated the subjects to speak. The graduated, meaningful, individual oral practice provided by the multimedia package seemed to help the experimental subjects overcome the "communication apprehension" (Curran, 1976; Friedman, 1980) that students normally endure trying to express themselves in the foreign language, and at the same time being reduced to a linguistically infantile level.

**Implications**

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 potential implications of subtitles and multimedia courseware for speaking practice in second/foreign language teaching and learning.

As concerns the subtitles, four implications are in order. 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 their potential effectiveness to the early stages of language learning, and to either visually- or aurally-oriented learners. Second, against arguments which contend that early introduction of writing may be negative, and in favor of positions which argue for the promotion of formal accuracy of learner output (Leech, 1974; Rivers, 1991), subtitles could make both the comprehension of "authentic" input and the production of "accurate" oral/written output less "painful." Third, developing learners' "taste" of subtitles would increase learners' interest in maintaining and/or increasing their knowledge of the foreign language outside the academic setting -- 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.

Referring to multimedia courseware, the main implication of this study is that future investigations on the effectiveness of computerized speaking practice, and CALL in general, should be based on made-to-measure, well-designed courseware. Indeed, the inability of computers to address language skills, and specially speaking skills, resides not on the computers themselves, as some would say (Ariew, 1987; Meunier-Cinko, 1992), but rather on the lack of courseware of an appropriate quality. To fill this void, design efforts of CALL investigators should be directed at: 1) developing courseware narrowly focused on the skill(s) to be taught and assessed; 2) ensuring that a sufficient amount of exposure to the experimental variable(s) is built into the research design in order to warrant the detection of potential experimental effects, and 3) selecting the appropriate learning tasks to engage experimental subjects in contextualized, meaningful practice.

If CALL is to be a valid alternative/supplement to traditional language learning, more task-oriented courseware has to be created and tested under conditions as close as possible to the real classroom. This advice is all the
more important as a surge of sophisticated but frequently "empty"
courseware products, often representing corporate interests, is invading the
educational arena at the risk of "trivializing" learning.

**Recommendations for Future Research**

In designing this study, the researcher drew mainly from theoretical
knowledge since there was little or no empirical information available. She
took into account theoretical recommendations regarding the choice of
independent variables, subtitling, (Salomon, 1979) and task level (Salomon,
1981). She also paid attention to the advice of designers concerning the
selection of descriptive measures by incorporating some of those that
according to Marchionini (1988) may advance the empirical knowledge
about hypermedia learning:

> If our goals in providing hypermedia assignments are related to processes and interactions, then we
must invent new strategies of evaluation that
address interactions. Both quantitative (i.e., time,
number of nodes connected, number of key
paths discovered) and qualitative (e.g.,
appropriateness of path, satisfaction of
experience) measures must be used if we are to
gain true images of how students are augmenting
their intellect with hypermedia. (p. 12)

Since the findings of this study support the above theories, it is
possible that a new empirical path had been open. The researcher would,
therefore, recommend to future investigators of this issue to follow the path by engaging in any of the following inquiries.

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.

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.

The study has proven that giving the learner 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 or not the additional possibility of inserting and removing the subtitles may increase such benefits.

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 and 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.

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, as mentioned earlier, to equipment limitations, text color coding has successfully been used in various microcomputer programs (Meara, 1985).

Another avenue of research that may be useful includes treating video segment length as a variable. In this study, the length of the subtitled video segments was based on theorist's recommendations about the optimal length of unsubtitled video to maintain learner's attention. It would be interesting to determine to which point the learner's attention and use of "chunking" strategies (Vanderplank, 1988) vary as the length of subtitled segments increases.

Finally, the investigation of how the organization of the speaking task may affect its "relevancy" constitutes an especially promising area for future research. Because students' acquisition of usefully organized knowledge cannot be taken for granted, much more must be learned about how knowledge organization models which have been successfully used in the sciences can be applied to multimedia programs for speaking practice.
If it is to be shown that it is "the application of the technology of education rather than the provision of technology in education" (Spencer, 1991: 21) that affects learning performance, then the above recommendations have to be put into practice.

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 those studies. The challenge of such efforts is to identify the appropriate speaking practice to be presented on the appropriate form to the appropriate students. The opportunity is the improvement of computerized language learning and education in general.
Notes to Chapter 1

(1) The Language Instruction through Communication-based Software (LINCS), created by Murray, Douglas, and Furstenberg (1989) within the context of the Athena Language Learning Project at MIT, has partially solved the problem of computers' inability to "speak." Communication in this software takes place in a written form: students type their messages and the computer, which has been designed with a maximal tolerance of errors, detects and answers them in a face-to-face conversation way.

(2) The practical requirement of a task-based test of oral performance is that elicitation of output criteria be constant. The testees have to perform an oral task using the target language, but they are not asked to use any specific grammatical structure or vocabulary in their performances.

(3) As used in this study, "locus of control" refers to the notion of student control versus program control of a computerized learning task. The concept of student control in CALL does not necessarily mean the type of internal control referred to in social learning theory (Bandura, 1977). According to this theory, internal control develops if certain conditions exist which enable learners gradually to develop competencies. In CALL, internal control is often akin to being put in charge of one's learning even if, as Robinson (1991) observes, "being put in charge and being in control are not synonymous." (p. 158).

(4) It is my opinion that Salomon's use of the word "short-circuiting" refers to the word "short cutting."

(5) Interaction with a medium code leads to its internalization just as, according to language theories (Vygotsky, 1962), interaction with the language community leads to the acquisition of the internal speech. While agreeing with this view, Salomon denied its feasibility and put observational learning as the cause of a code's internalization. Salomon's denial is understandable since the medium upon which he built his theories, TV, does not allow the kind of interaction offered by new delivery systems such as IVD.
(6) According to Meredith (1983) both studies should be taken as examples of good basic research. The studies were designed to determine how variations in the use of a specific medium affect language learning.

(7) Swain (1985) contends that only meaning-focused exchanges may prompt learners to shift from semantic to syntactic processing, thus producing meaningful output. Contrary to Swain's view, substantial research (Long, 1983; Ellis, 1985) has proven that meaningful output can also be facilitated by formal instruction -- such as the one offered by the courseware package used in this study.

Notes to Chapter 2

(1) Brown and Yule (1983b) 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 (e.g.: justifying a course of action) relationships.

(2) From a linguistic point of view, and following Jung's (1990) terminology, there are two kinds of subtitles: "intralingual" and "interlingual." 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". By this is meant the possibility of completely translating the spoken dialogue/presenter or voice-over commentary.

From a technological point of view, a distinction can be made between subtitled video (subtitles inserted on a videotape or laser disc), and broadcast videotext, or closed-captioned video (subtitles inserted on four "empty" lines of an ordinary 625-line TV picture which can be brought on and off at will).

(3) A decoder is a device which when attached to a TV can "capture" a coded signal. The decoder places narrow black boxes onto the TV screen. The boxes contain the closed-captions of the program's dialogue.

(4) At least three approaches to the communicative competence concept appear to have influenced the definition of the concept: 1) the psychological and communication approach (Wang, Rose, & Maxwell,
1973) which defines communicative performance as the ability to receive and convey information; 2) the sociolinguistic approach (Hymes, 1972) which stresses the appropriateness of communication according to parameters such as topic, place, or interlocutor; and 3) the speech act analysis approach (Searle, 1969) which emphasizes the effective performance of speech acts as the key to communicative competence.

(5) According to Clark (1979), direct tests have a high degree of face/content validity but are expensive to administer and are not particularly efficient in diagnosing specific weaknesses. On the other hand, semi-direct tests permit the systematic presentation of highly focused stimuli in a way that is not generally feasible in a live conversational setting. The major disadvantage of semi-direct tests is that they may have varying degrees of scoring reliability.

Notes to Chapter 3

(1) The Proficiency Guidelines were developed in 1981, and modified in 1986, by the American Council on the Teaching of Foreign Languages (ACTFL). Modelled on the rating descriptions used in the oral interview, these Guidelines address speaking, listening, reading, and writing. They were designed to guide curriculum and materials development as well as to provide a graduated sequence of learning goals for teachers and students.

(2) The term "performance" as used in this study relates to acquisition of specific linguistic features that are associated with a particular course of instruction. It contrast with the term "proficiency" which refers to an acquired competence in the language regardless of how or where the language was acquired.

(3) To avoid pretest sensitization (Borg & Gall, 1989), no assessment of the subjects' ability to describe or narrate was made before the experimental sessions.

(4) The oral communicative competence construct draws on Spolsky's Preference Model (1989). From a practical point of view, in the development of language tests, this model is less restrictive than competing models such as Canale and Swain's (1980), or its modified version Canale's (1983). Indeed, whereas application of Canale's model supposes the a priori
inclusion of its four competence components (grammatical, sociolinguistic, discourse, and strategic), application of Spolsky’s model allows for the setting of possible dimensions, the planned uses of the test determining what importance should be given to measuring each of the dimensions.

(5) The random oral samples for inter-rater reliability assessment were taken from the first of the experimental practice sessions, instead of from the pilot study, for two reasons: first, it was thought that scores on the pilot study’s oral samples were not completely representative of scores of the experimental study’s samples because a number of pilot subjects had a higher language level, French 4100, than the experimental subjects; second, since the first practice session was conceived as a training session, it was also thought that scores obtained in this session could rightly be used for piloting purposes.

(6) Given the small sample size of the pilot study, only questionnaire form B, the one which could be answered by subjects in the four treatments, was administered for validation to all subjects.

Notes to Chapter 4

(1) Format which enables the user to browse through information that is linked, cross-referenced and annotated.

(2) Users can either type out in full or click the answer of their choice. To avoid any cheating, answers are automatically locked when users click on the "See scores" button. All the possible correct answers have been precoded and recorded by alphanumeric codes to provide users with immediate feedback. Since data entry and error checking are concurrent, the program can produce a fully cleaned data file ready for analysis shortly after the completion of field work.

(3) Somewhat simplified, the flowchart doesn’t show all the options available in the "Très Utile" stack.

(4) These learning processes serve the following functions: 1) alerting to receive stimulation, 2) acquiring an expectancy of the results of learning, 3) retrieving items from the long term memory to the working memory, 4) selecting learning patterns, 5) encoding presented material, 6) responding
with a performance that verifies learning, 6) reinforcing, 7) cueing, and 8) generalizing performance to new situations.

[5] The materials represent and depict the following video genres and functional speech act situations: 1) documentary feature (monologic/dialogic speech with focus on narrative/descriptive details, reported facts and events); 2) commercials (monologic speech with emphasis on register, emotional tone, situational humor or irony, and added comprehension layer of repetition or rhythm); 3) video clips (lyric speech with added elements of repetition, rhyme and rhythm).

[6] The works of Garza (1986), Joiner (1990), and Grimes (1990) were taken into account in defining the quality and difficulty criteria for segments selection.

[7] Three out of the four selected segments were kept entirely intact as they appeared in the original source; the remaining segment was internally edited because it was too long.

[8] In contrast with the usual closed-captioning technique (paraphrased or simplified verbatim of the original audio script) used by the National Captioning Institute (NCI), this study has employed the open-captioning technique (verbatim of the experimental video segments) utilized in recent Second Language captioning research (Garza 1991).

[9] Attempts which take into account designers's recommendations (Keller & Suzuki, 1988) and principles applied by sound CALL programs (Robinson et al., 1985).

[10] That is, by using bolded titles, words and phrases that cue idea importance and connections between levels of the explicit arguments.

Notes to Chapter 5

[1] Figure 5. 16 in previous chapter has been created with Aldus FreeHand 2.02 (1988). Figures in this chapter have been created with Cricket Graph 1.2 (1988) or generated by the SuperANOVA (1991) statistics package.
Eight out of the ten sets of data were displayed on the "Control Card" and the remaining ones (length of recording and number of trials) on the "Recording Studio."

Notes to Chapter 6

(1) There is evidence that under a learner-control condition, students tend to spend less time on the lesson than subjects in a learner-control with advisement condition (Johansen & Tennyson, 1983).

(2) Times of the subtitles subjects were supposed to be longer than times of the no-subtitles subjects since the former had to pay attention to the information coming from both the visual and the aural channels.

(3) In a recent article about non-native speakers use of teletex subtitles in English, Vanderplank (1992) comments on the "good" language ability, and the positive attitude of the general population of the Nordic countries and Holland to viewing foreign language programs which are invariably accompanied by interlingual subtitles.

(4) Descriptive data indicated that subtitles subjects entered "Très Utile" more frequently and stayed there longer than no-subtitles subjects. This was an interesting finding because it was thought that the former would explore "Très Utile" less than the latter since they had an additional source of information, subtitles, available on the main path.

(5) Channel layer density, or sensory overload, is frequently attributed (Reese, 1982) to the inability of television to allow for the rehearsal of information.

(6) Unfortunately, as Spencer (1991) points out, provision of technology "is frequently given priority when non-educational agencies are involved, ill-informed decision-makers being seduced by the superficial appeal of new hardware approaches." (p. 21).
REFERENCES


212


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.


Meeting of the Association of Developers of Computer Based
Instructional Systems, Dallas, TX. (ERIC Document Reproduction
Service No. ED 165 716).

Lambert, W. E., Boehler, I., & Sidoti, N. (1981). Choosing the languages of
subtitles and spoken dialogues for media presentations: Implications
for second language education. *Applied Psycholinguistics, 2*, 133-
148.

Publications, Ltd.


Adult Literacy & basic skills Unit, ILEA.


Likert, R. (1932). Techniques for the Measurement of Attitudes. *Archives of
Psychology, 140*, 5-55.

Council of Teachers of English.


Lowe, P., Jr., & Clifford, R. T. (1980). Developing an indirect measure of
overall proficiency. In J. R. Frith (Ed.), *Measuring spoken language
proficiency* (pp. 31-39). Washington D.C.: Georgetown University
Press.

*Colorado CAI Spanish Research Project*. Final Report. Boulder, CO:
University of Colorado, International Research and Studies Program.

M. C. Wittrock (Ed.), *The brain and psychology* (pp. 141-210). New


Oller, J. W., Jr. (1979). Explaining the reliable variance in test: the validation problem. In E. J. Brière, & F. B. Hinofotis (Eds.), *Concepts*
in language testing: Some recent studies (pp. 61-74). Washington, DC: TESOL.


*SPSS* [software]. (1986). Chicago, IL: SPSS International B. V.


Appendix A

Invitation Letter to Experimental Subjects

Isabel Borrás
French Education Project
Louisiana State University
202 Peabody Hall
Baton Rouge, LA 70803

April 7, 1992

Dear X:

You have been selected by the College of Education to participate in two French practice sessions designed to help researchers to assess the effectiveness of multimedia materials in foreign language teaching.

In exchange for your attendance at these sessions, you will be excused from two French classes.

Check your name on the scheduling lists that will be distributed by your instructor, and sign up for the day and time of your convenience.

Please contact me if you have any question. I can be reached at home (504) 768-7344 or at work (504) 388-2309.

Thank you very much.

Cordially,

Isabel Borrás

232
Appendix B

Oral Communicative Performance: Evaluation Form

RECORDING FORM

Directions are given for scoring each of the tasks accomplished by the subjects during the experimental sessions, and space is provided for recording performance. The general criterion for adequate performance on all tasks is a straightforward answer to this question: Has the subject addressed the nature of the task in an appropriate manner? Transcribe recorded oral samples verbatim on the provided spaces. Record scores as ratios of scales' highest level (e.g.: 4/6), and weigh by 3 the scores obtained in the Effectiveness scale. After reviewing performance on all tasks, complete the Performance Summary.

This recording form shouldn't be completed until you have thoroughly viewed the video segments, read the video segments' transcripts, and clearly understood the purpose of each task.

Tasks: Session 1

Task 1: "Pub" (Description)

Purpose of task: To selectively feature and subordinate details and use them to describe a person.

Directions for the student:

- Using the computer keyboard, or pencil and paper, draft a description of Luc Châtelain, the French advertising "star" portrayed in the video segment "Les enfants et la publicité."

- Your description has to include the following: 1) one of the two settings where he appears inside his house (kitchen, study room); 2) his identity (nationality, profession, father's profession; 3) his physical appearance (age, body complexion, hair, face, etc.); 4) his clothes; 5) his character.

Directions for the evaluator:
(Appendix B continued)

Performance: (Transcribe what you understand from recorded description)

Scoring Instructions: Rate recorded description on each of the four scales on page 3 of this form.

Scales: Effectiveness  Accuracy  Organization  Fluency

Scores: ___/___  ___/___  ___/___  ___/___  ___/___

Total: __________

Comments:

Task 2: "Echecs" (Narration)

Purpose of task: To use cognitive organization to systematically gather information and formulate a coherent narrative.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Directions for the student:

- Using the computer keyboard, or pencil and paper, draft a narration of the documentary you have seen, "Le jeune champion d'échecs."
- You have to provide some details about the documentary's main character (eg: Adrien's age and profession), and about some of the documentary's locations (eg: Adrien's house, bookstore, streets); 2) report at least five of the things Adrien does or says throughout the video segment.
- Recount your story in the present tense and be consistent with the use of this verb time throughout your narration.

Directions for the evaluator:

Performance: (Transcribe what you understand from recorded narration)

Scoring Instructions: Rate recorded narration on each of the four scales on page 3 of this form.
Tasks: Session 2

Task 1: "Village" (Description)

Purpose of task: To selectively feature and subordinate details and use them to describe a person.

Directions for the student:

Using the computer keyboard, or pencil and paper, draft a description of Paul Vouillard, the main character of the video segment "Reconstruction d'un village."

Your description has to include the following: 1) one of the places where M. Vouillard appears in the village (eg: street, square, hill); 2) his identity (former profession and town where he practiced it); 3) his physical appearance (estimated age, body complexion, hair, face, etc.); 4) his clothes; 5) his character.

Directions for the evaluator:
Performance: (Transcribe what you understand from recorded description)

Scoring Instructions: Rate recorded description on each of the four scales on page 3 of this form.

Scales: Effectiveness Accuracy Organization Fluency

Scores: ___ / ___ ___ / ___ ___ / ___ ___ / ___

Total: _________

Comments:

Task 2: "Bac" (Narration)

Purpose of task: To use cognitive organization to systematically gather information and formulate a coherent narrative.

Directions for the student:

. Using the computer keyboard, or pencil and paper, draft a narration of the documentary you have seen, "Le baccalauréat."

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
(Appendix B continued)

. Your have to: 1) provide some details about the documentary's main character (eg: Chrystèle Bremard's age and studies), and about some of the documentary's locations (eg: Chrystèle and friend's apartments, department store); 2) report at least five of the things Chrystèle does or says throughout the video segment.

. Recount your story in the present tense and be consistent with the use of this verb time throughout your narration.

Directions for the evaluator:

Performance: (Transcribe what you understand from recorded narration)

Scoring Instructions: Rate recorded narration on each of the four scales on page 3 of this form.

Scales: Effectiveness Accuracy Organization Fluency

Scores: ___/___  ___/___  ___/___  ___/___

Total: _________

Comments:
RATING SCALES

Recorded oral samples are to be rated on the below six-point scales. The scales are to reflect subjects oral performance at two levels: 1) global (Effectiveness scale), and 2) component (Accuracy, Organization, and Fluency scales).

EFFECTIVENESS SCALE (adapted from Bartz, 1979)

General definition: Amount of relevant information conveyed by the subject.

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 means "relevant information" 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.
(Appendix B continued)

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 "Narration" samples, evaluators 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 inter-sentential 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 hesitancies.

5. Utterances quite fast and fairly effortless.

6. Utterances produced with a native speaker's speed and ease.

NOTE: When applying this scale, evaluators should annotate the length of the recordings.
PERFORMANCE SUMMARY

Student: .........................................................................................................................

ID n° .................................................................................................................................

Course: .................................. Section: ................................................................. Date: ..............

Evaluator: ....................................................................................................................... Position: ........................................

<table>
<thead>
<tr>
<th>TASKS</th>
<th>Pubs (Training/Description)</th>
<th>&quot;Echecs&quot; (Training/Narration)</th>
<th>&quot;Village&quot; (Evaluation/Description)</th>
<th>&quot;Bac&quot; (Evaluation/Narration)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCALES</td>
<td>Effectiveness</td>
<td>Accuracy</td>
<td>Organization</td>
<td>Fluency</td>
</tr>
<tr>
<td></td>
<td><em><strong>/</strong></em></td>
<td><em><strong>/</strong></em></td>
<td><em><strong>/</strong></em></td>
<td><em><strong>/</strong></em></td>
</tr>
<tr>
<td></td>
<td><em><strong>/</strong></em></td>
<td><em><strong>/</strong></em></td>
<td><em><strong>/</strong></em></td>
<td><em><strong>/</strong></em></td>
</tr>
<tr>
<td></td>
<td><em><strong>/</strong></em></td>
<td><em><strong>/</strong></em></td>
<td><em><strong>/</strong></em></td>
<td><em><strong>/</strong></em></td>
</tr>
<tr>
<td></td>
<td><em><strong>/</strong></em></td>
<td><em><strong>/</strong></em></td>
<td><em><strong>/</strong></em></td>
<td><em><strong>/</strong></em></td>
</tr>
</tbody>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
APPENDIX C

Subjects' Oral Samples: Eight Transcripts

PROGRAM 1

Sample: 1. Task: Description. Treatment: T1 (Subtitles/Lower-Level Task)

"Les enfants et la publicité est sur un petit garçon qui s'appelle Luc Châtelain. Il a douze ans. Il a commencé à travailler dans la publicité quand il avait deux ans et demi. Il a les cheveux roux et les yeux bruns. Son visage est rond et... il a des taches de rousseur. Il a des lèvres fines. Il est de petite taille. Il a l'air très marrant. Il parle un peu mais il sourit beaucoup. Dans l'appartement de ses parents, il répond aux questions d'une journaliste... Il porte un T-shirt blanc et des shorts. Son père est boucher. Sa mère travaille à demi-temps. Elle l'accompagne aux plateaux de tournage et aux prises de vue."

Sample: 2. Task: Description. Treatment: T2 (No-Subtitles/Lower-Level Task)

"Il s'appelle Luc Châtelain. Il est français. Il a douze ans. Son père est un boucher. Luc est très petit et mince. Il a les cheveux en brosse. Il a le visage ovale... Il a les yeux bleus. Il est le star de quarante films publicitaires. Il est sympa et il a l'air gentil... Quand il est dans la cuisine il porte un T-shirt et un jean."

Sample: 3. Task: Description. Treatment: T3 (Subtitles/Higher-Level Task)

"C'est un program au sujet des enfants qui travaillent. En France, c'est interdit pour l'enfant de travailler... mais il y a des exceptions: dans les publicités ou des films, il y a des enfants qui travaillent. Un exemple est un garçon qui s'appelle Luc Châtelain. Il a douze ans et il a commencé à travailler à l'âge de deux ans et demi... Il a les cheveux roux, en brosse. Il a de
grands yeux et des tâches de rousseur sur les joues. Il gagne beaucoup d'argent mais sa mère ne veut pas dire combien son fils gagne. Son père est boucher. Il porte un T-shirt. Il mange des biscottes avec de la confiture dans la cuisine de sa maison quand il est interviewé. Il semble un enfant ordinaire et il sourit beaucoup en disant que l'argent ne lui pèse pas lourd."

Sample: 4. Task: Description. Treatment: T4 (No-Subtitles/Higher-Level Task)

"Luc Châtelain a commencé son travail à un âge très jeune. Il avait deux ans et demi quand il avait son début. Maintenant il a douze ans et il est une star. Il a les cheveux roux et il a aussi les cheveux en brosse. Il est un très heureux enfant... Dans ses annonces, il porte beaucoup de choses différentes. Il porte des vêtements d'hiver, un T-shirt, des pullovers de beaucoup de couleurs différents... de pantalons et ainsi de suite. Si on le voit dans sa chambre, il porte un T-shirt et des bermudas. Il est assis à son bureau où il fait ses devoirs. Il a beaucoup de livres avec des photos. Et... au-dessus de son bureau il a une grande carte du monde."

PROGRAM 2

Sample: 5. Task: Narration. Treatment: T1 (Subtitles/Lower-Level Task)

"C'est le mois de juin en France et il y a beaucoup de jeunes gens qui doivent étudier pour le bac --c'est-à-dire, le baccalauréat. On trouve une jeune fille qui s'appelle Chrystèle. Elle étudie long et dur pour le bac. Elle étudie pour la plupart la philo et les mathématiques. Le rite du jour pour Chrystèle sont comme ça: d'abord elle fait du jogging pour se vider l'esprit. Puis elle rentre chez elle et elle commence ses études... Elle a besoin de bonbons pour éviter le trac. Puis elle sort chez son amie qui..."
regarde un match de tennis. Après..., Chrystèle sort de la maison de son amie et rentre chez elle... Elle y recommence ses études. Elle a moins que vingt-quatre heures pour passer son bac. En bref, elle dit que pour une bonne révision il faut de bons outils, des annales, des cours, de petits bouquins,... des résumés --et, bien sûr, des bonbons! A la fin de l'émission elle retourne aux études et elle planche pour le bac."

Sample: 6. Task: Narration. Treatment: T2 (No-Subtitles/Lower-Level Task)


Sample: 7. Task: Narration. Treatment: T3 (Subtitles/Higher-Level Task)

"Le personnage central de cette petite histoire 'Le baccalauréat' est une fille, Chrystèle, qui a dix-huit ans... Demain pour elle, c'est le bac, l'examen le plus difficile pour les jeunes gens français. Cette année passée au lycée, Chrystèle a étudié la philo et les maths... c'est ce que nous la voyons en train de réviser. Quand nous la voyons au début, elle fait du jogging --d'après le narrateur, 'pour se vider l'esprit'. Après le sport, depuis deux semaines, elle révise au moins cinq par jour. En bref, elle explique ce qu'il faut faire pour bien réviser. Elle dit que la chose la plus importante, c'est ce que le prof a dit dans l'année... Plus tard, elle va au supermarché pour acheter des bonbons. Alors, elle les prend chez sa meilleure copine. Là, les deux se rassurent l'une l'autre. C'est déjà quinze heures mais ni Chrystèle ni son amie semblent avoir peur. Après sa visite, Chrystèle rentre chez
elle pour réviser les maths. Elle étudie un peu plus. À vingt et une heures, c’est le temps pour dormir. Le narrateur n’est pas certain que ce soit possible car elle est nerveuse.

Sample: 8. Task: Narration. Treatment: T4 (No-Subtitles/Higher-Level Task)

“Chrystèle Bremard est une jeune lycéenne française qui prépare son bac. Elle a dix huit ans et elle habite dans une ville en France. Chrystèle étudie cinq ou six heures par jour pour les classes de math et philosophie. Pour bien réviser, elle a besoin de bons outils comme un livre et un bouquin... Chrystèle fait du jogging pour sa bonne santé... Plus tard, elle visite sa copine qui regarde une match de tennis à la télévision. Pour éviter le trac, elle mange des bonbons -- peut-être, c’est pourquoi elle fait du jogging. Après avoir visité sa copine, Chrystèle rentre chez elle pour réviser les maths. Finalement elle fait dodo.”
APPENDIX D

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."

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

247
6. The questions and the feedback of the trivial pursuit and quizzes helped me to better understand and remember the video segments' content.

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

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.

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.

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

Attitude Questionnaire Form B

The information you provide in this questionnaire will be of great value for the designer of the multimedia package "Practicing Spoken French."

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

249
6. The questions and the feedback of the trivial pursuit and quizzes helped me to better understand and remember the video segments' content.

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

8. I found very useful the possibility of controlling the videodisc player from the computer screen.

9. The high quality of the video images played an important role in understanding the video information.

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

"Practicing Spoken French": Instructions to the User

To start your practice session:

1. Press the ^ key at the upper right corner of the keyboard.
2. Double-click on the "Macintosh HD" icon at the upper right of the screen.
3. Double-click on the folder "Isabel" inside the hard drive.
4. Double-click on the stack of your name.

To conclude your practice session:

1. Click on the END button of the Home Card of the program you are working on.
2. Close the window "Isabel" by clicking on the window's upper left corner.
3. Close the window "Macintosh HD" by clicking on the window's upper left corner.
4. Select "Special" from the menu bar and click on "Shut Down".

GOOD LUCK!

Isabel Borràs
(Appendix F continued)

How to Get French & Spanish Characters on the Macintosh U.S. Keyboard

à (Option+’)+a  á (Option+e)+a  â (Option+i)+a  ā (Option+u)+a

e (Option+’)+e  é (Option+e)+e  ê (Option+i)+e  ē (Option+u)+e

ï (Option+i)+i  ï (Option+u)+i

ô (Option+i)+o  ò (Option+u)+o

û (Option+’)+u  û (Option+e)+u  ü (Option+i)+u  ü (Option+u)+u

Œ (Option+n)+n  Ç (Option+c)  Ç (Shift+Option+c)

¿ (Option+?)  I (Option+1)

* Commands within parentheses have to be pressed together.
Appendix G

Courseware Development: Staff Hours

Below there is a "conservative" estimate of the tasks and number of staff hours spent in the design and program of "Practicing Spoken French." The estimate does not reflect the time invested by two post-production companies on the mastering and pressing of the videodisc segments.

<table>
<thead>
<tr>
<th>Staff</th>
<th>Tasks</th>
<th>Tasks' specification</th>
<th>Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>This researcher</td>
<td>Design of two programs with two instructional units each.</td>
<td>Select and transcribe video segments.</td>
<td>300</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Design units' framework (narrative learner-system conversation).</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Design units' storyboard.</td>
<td></td>
</tr>
<tr>
<td>This researcher + five</td>
<td>Computer-recording of units' introductory messages and useful expressions.</td>
<td>Rehearse Record Transfer</td>
<td>15</td>
</tr>
<tr>
<td>native speakers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(French, English)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

253
This researcher + one programmer

Programming of units

Create HyperCard stacks.

Design animations.

Design scoring systems of listening comprehension.

Import native speakers' voices and music excerpts.

Scan and import pictures.

Digitalize video images.

Generate video subtitles.
Appendix H

PSF Documentation

"Practicing Spoken French" builds upon the following materials:

Animations: Five animated texts (drag effect).

Graphics: Nine scanned pictures, one of them portraying a group of faculty members at LSU, and the others representing two Louisiana settings: the "University Lake" (Baton Rouge), and the "City Park" (New Orleans).

Sounds:
- Recorded voices: native speakers and users.
- Music excerpts:
  -- "Arabesque" (Rondo Veneziano)
  -- "Corso Delle Gondole" (Rondo Veneziano)
  -- "Bourré" (Händel)
  -- "La Scala d'Oro" (Rondo Veneziano)
  -- "La Réjouissance" (Händel)
- Other music/sounds:
  -- "LogoLick"
  -- "Horn Fanfare"
  -- Footsteps

Video: Eight video segments (four with subtitles and four without) taken from the French video series France Panorama and pressed onto videodisc. The segments include:
  -- "Les enfants et la publicité" (02:35)
  -- "Le jeune champion d'échecs" (02:16)
  -- "Reconstruction d'un village" (02:09)
  -- "Le baccalauréat" (02:30)

255

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
### APPENDIX I

**Preselected Video Segments**

<table>
<thead>
<tr>
<th>Title</th>
<th>Tape N°</th>
<th>Ref. on tape</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Segment 1  Le Bac (1ʳᵉ partie)</td>
<td>III/1</td>
<td>09:04</td>
<td>02:30</td>
</tr>
<tr>
<td>Segment 2  Stylo pour aveugles</td>
<td>III/1</td>
<td>29:23</td>
<td>01:45</td>
</tr>
<tr>
<td>Segment 3  IGN</td>
<td>III/1</td>
<td>37:50</td>
<td>01:26</td>
</tr>
<tr>
<td>Segment 4  Océanopolis</td>
<td>III/3</td>
<td>05:02</td>
<td>01:44</td>
</tr>
<tr>
<td>Segment 5  Le boulanger de Sardent</td>
<td>III/3</td>
<td>19:34</td>
<td>02:03</td>
</tr>
<tr>
<td>Segment 6  Le Petit Larousse et les amanites phalloïdes</td>
<td>III/4</td>
<td>31:23</td>
<td>01:34</td>
</tr>
<tr>
<td>Segment 7  Les enfants et la pub (Luc Châtelain)</td>
<td>III/5</td>
<td>35:21-35:55</td>
<td>02:35</td>
</tr>
<tr>
<td>Segment 8  Les renifleurs</td>
<td>III/6</td>
<td>38:40</td>
<td>2:19</td>
</tr>
<tr>
<td>Segment 9  Les femmes et la politique</td>
<td>IV/1</td>
<td>04:52</td>
<td>02:00</td>
</tr>
<tr>
<td>Segment 10 Un champion d’échecs de dix ans</td>
<td>IV/1</td>
<td>20:05</td>
<td>02:16</td>
</tr>
</tbody>
</table>

256

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
(Appendix I continued)

<table>
<thead>
<tr>
<th>Segment</th>
<th>Title</th>
<th>Channel</th>
<th>Duration</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>Les Vosges et l'environnement</td>
<td>IV/1</td>
<td>29:30</td>
<td>02:07</td>
</tr>
<tr>
<td>12</td>
<td>Petite fille alpiniste</td>
<td>IV/2</td>
<td>00:30</td>
<td>02:21</td>
</tr>
<tr>
<td>13</td>
<td>Vieille dame</td>
<td>IV/2</td>
<td>24:44</td>
<td>01:52</td>
</tr>
<tr>
<td>14</td>
<td>Les cigognes en Alsace</td>
<td>IV/2</td>
<td>26:37</td>
<td>02:19</td>
</tr>
<tr>
<td>15</td>
<td>L'Imprimerie Nationale</td>
<td>IV/2</td>
<td>32:40</td>
<td>03:17</td>
</tr>
<tr>
<td>16</td>
<td>Reconstruction d'un village</td>
<td>IV/2</td>
<td>35:58</td>
<td>02:09</td>
</tr>
</tbody>
</table>
APPENDIX J

Instrument for Rating the Quality and Difficulty of the Experimental Video Segments

The 16 video segments you are about to see are being considered for use in a speaking experiment with 60 fifth semester college students of French. Please look at the grid below and rate on a scale of 1 to 5 each of the video segments listed on the left-hand column according to the criteria on the upper row. Any comments you might have about the reasons for your ratings would be greatly appreciated. (Rating values and criteria are defined on the next page of this instrument).

<table>
<thead>
<tr>
<th>Scales</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Segment 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Segment 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Segment 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Segment 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Segment 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Segment 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Segment 7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Segment 8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Segment 9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Segment 10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Segment 11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Segment 12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Segment 13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Segment 14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Segment 15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Segment 16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Rating values:

Quality: 1 (very poor), 2 (poor), 3 (fair), 4 (good), 5 (excellent).

Difficulty: 1 (very easy), 2 (easy), 3 (average), 4 (difficult), 5 (very difficult).

Rating criteria:

Quality:

. Situational appropriateness: The language is appropriate for the situation portrayed in the video segment.

. Inherent interest value to university level students: The subject matter of the segments is relevant and of interest to the average fifth semester university student.

. Audio/visual correlation: The language of the audio track is supported and contextualized by the images of the segment.

Difficulty:

. Grammatical and lexical complexity: Most of the language in the segment is appropriate for fifth semester college students of French.

Please return your ratings as soon as possible. And thank you for sharing your time and expertise.
Appendix K

Selected Video Segments: Mean Ratings and Length

Session 1

<table>
<thead>
<tr>
<th>Segment</th>
<th>Quality mean</th>
<th>Difficulty mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. &quot;Enfants et pub&quot; (Description)</td>
<td>4.08</td>
<td>3.00</td>
</tr>
<tr>
<td>2. &quot;Champion d'échecs&quot; (Narration)</td>
<td>4.08</td>
<td>2.75</td>
</tr>
<tr>
<td>Overall means</td>
<td>4.08</td>
<td>2.87</td>
</tr>
</tbody>
</table>

Session 2

<table>
<thead>
<tr>
<th>Segment</th>
<th>Quality mean</th>
<th>Difficulty mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>3. &quot;Reconstruction d'un village&quot; (Description)</td>
<td>4.08</td>
<td>3.00</td>
</tr>
<tr>
<td>4. &quot;Bac&quot; (Narration)</td>
<td>4.16</td>
<td>3.00</td>
</tr>
<tr>
<td>Overall means</td>
<td>4.12</td>
<td>3.00</td>
</tr>
</tbody>
</table>

Segments' length: 1) 02:35; 2) 02:16; 3) 02:09; 4) 02:30.
Total segments' length: 09:30.
APPENDIX L

Video Segments Production: Who Did What?

The below companies and individuals, and the services they provided (specified in parentheses), made possible the production of the experimental video segments.

Eagle Multimedia Services Inc.
91 5th Avenue, Suite 800
New York, NY 10003
(Source video segments)

Edge Productions
101 Beaver Dam Rd.
Aiken, SC 20801
(Segments recording on SMPTE 1”)

Video Park
11316 Pennywood Av.
Baton Rouge, LA
(Segments transfer from 1” to SVHS)

Mr. Alexander Perlis
Louisiana State University
(Subtitling)

Magno Sound & Video
729 7th Avenue
New York, NY 10019
(Videodisc mastering)
Appendix M
Interaction Bar Charts of Oral Performance Criteria per Treatment and Program

Figure M.1 Interaction Bar Chart of Overall Oral Performance per Treatment and Program
Figure M. 2 Interaction Bar Chart of Oral Performance (Effectiveness) per Treatment and Program
Figure M. 3 Interaction Bar Chart of Oral Performance (Accuracy) per Treatment and Program
Figure M. 4 Interaction Bar Chart of Oral Performance (Organization) per Treatment and Program
Figure M.5 Interaction Bar Chart of Oral Performance (fluency) per Treatment and Program
Appendix N

Means and Standard Deviations of Overall Oral Performance Scores and Subscores (P1) as a Function of Subtitling and Task Level

Table N. 1  Means and Standard Deviations of Overall Performance Scores (P1) as a Function of Subtitling and Task Level

<table>
<thead>
<tr>
<th>Subtitles</th>
<th>Lower-level Task</th>
<th>Higher-level Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>Std. Dev.</td>
<td>Mean</td>
</tr>
<tr>
<td>Subtitles</td>
<td>32.25</td>
<td>1.815</td>
</tr>
<tr>
<td>No-Subtitles</td>
<td>27.10</td>
<td>2.025</td>
</tr>
</tbody>
</table>

Table N. 2  Means and Standard Deviations of Effectiveness Scores (P1) as a Function of Subtitling and Task Level

<table>
<thead>
<tr>
<th>Subtitles</th>
<th>Lower-level Task</th>
<th>Higher-level Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>Std. Dev.</td>
<td>Mean</td>
</tr>
<tr>
<td>Subtitles</td>
<td>17.50</td>
<td>1.000</td>
</tr>
<tr>
<td>No-Subtitles</td>
<td>15.70</td>
<td>1.059</td>
</tr>
</tbody>
</table>

267

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
### Table N. 3  Means and Standard Deviations of Accuracy Scores (P1) as a Function of Subtitling and Task Level

<table>
<thead>
<tr>
<th>Subtitles</th>
<th>Lower-level Task</th>
<th>Higher-level Task</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Dev.</td>
</tr>
<tr>
<td>Subtitles</td>
<td>4.83</td>
<td>.577</td>
</tr>
<tr>
<td>No-Subtitles</td>
<td>3.20</td>
<td>.919</td>
</tr>
</tbody>
</table>

### Table N. 4  Means and Standard Deviations of Organization Scores (P1) as a Function of Subtitling and Task Level

<table>
<thead>
<tr>
<th>Subtitles</th>
<th>Lower-level Task</th>
<th>Higher-level Task</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Dev.</td>
</tr>
<tr>
<td>Subtitles</td>
<td>5.00</td>
<td>.739</td>
</tr>
<tr>
<td>No-Subtitles</td>
<td>4.20</td>
<td>.789</td>
</tr>
</tbody>
</table>
Table N. 5  Means and Standard Deviations of Fluency Scores (P1) as a Function of Subtitling and Task Level

<table>
<thead>
<tr>
<th></th>
<th>Lower-level Task</th>
<th>Higher-level Task</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Dev.</td>
</tr>
<tr>
<td><strong>Subtitles</strong></td>
<td>4.92</td>
<td>.515</td>
</tr>
<tr>
<td><strong>No-Subtitles</strong></td>
<td>3.80</td>
<td>.632</td>
</tr>
</tbody>
</table>
Appendix O

Means and Standard Deviations of Overall Oral Performance Scores and Subscores (P2) as a Function of Subtitling and Task Level

<table>
<thead>
<tr>
<th>Subtitles</th>
<th>Lower-Level Task</th>
<th>Higher-level Task</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Dev.</td>
</tr>
<tr>
<td>Subtitles</td>
<td>32.92</td>
<td>1.975</td>
</tr>
<tr>
<td>No-Subtitles</td>
<td>27.70</td>
<td>2.214</td>
</tr>
</tbody>
</table>

Table O. 2 Means and Standard Deviations of Effectiveness Scores (P2) as a Function of Subtitling and Task Level

<table>
<thead>
<tr>
<th>Subtitles</th>
<th>Lower-level Task</th>
<th>Higher-level Task</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Dev.</td>
</tr>
<tr>
<td>Subtitles</td>
<td>17.67</td>
<td>.651</td>
</tr>
<tr>
<td>No-Subtitles</td>
<td>15.30</td>
<td>1.252</td>
</tr>
</tbody>
</table>
### Table O. 3  Means and Standard Deviations of Accuracy Scores (P2) as a Function of Subtitling and Task Level

<table>
<thead>
<tr>
<th></th>
<th>Lower-level Task</th>
<th>Higher-level Task</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Dev.</td>
</tr>
<tr>
<td><strong>Subtitles</strong></td>
<td>4.92</td>
<td>.793</td>
</tr>
<tr>
<td><strong>No-Subtitles</strong></td>
<td>4.00</td>
<td>.943</td>
</tr>
</tbody>
</table>

### Table O. 4  Means and Standard Deviations of Organization Scores (P2) as a Function of Subtitling and Task Level

<table>
<thead>
<tr>
<th></th>
<th>Lower-level Task</th>
<th>Higher-level Task</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Dev.</td>
</tr>
<tr>
<td><strong>Subtitles</strong></td>
<td>5.25</td>
<td>.754</td>
</tr>
<tr>
<td><strong>No-Subtitles</strong></td>
<td>4.30</td>
<td>.823</td>
</tr>
</tbody>
</table>
Table O. 5  Means and Standard Deviations of Fluency Scores (P2) as a Function of Subtitling and Task Level

<table>
<thead>
<tr>
<th></th>
<th>Lower-level Task</th>
<th></th>
<th>Higher-level Task</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Dev.</td>
<td>Mean</td>
<td>Std. Dev.</td>
</tr>
<tr>
<td>Subtitles</td>
<td>5.08</td>
<td>.515</td>
<td>5.36</td>
<td>.505</td>
</tr>
<tr>
<td>No-Subtitles</td>
<td>4.10</td>
<td>.568</td>
<td>4.36</td>
<td>.674</td>
</tr>
</tbody>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Appendix P

Means and Standard Deviations of Overall Oral Performance Scores (P1, P2) as a Function of Subtitling and Task Type

Table P. 1  Means and Standard Deviations of Overall Oral Performance Scores (P1) as a Function of Subtitling and Task Type

<table>
<thead>
<tr>
<th>Description</th>
<th>Subtitles</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>No-subtitles</th>
<th>Mean</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Subtitles</td>
<td>32.333</td>
<td>1.775</td>
<td>27.667</td>
<td>2.640</td>
<td></td>
</tr>
<tr>
<td></td>
<td>No-subtitles</td>
<td>27.667</td>
<td>2.040</td>
<td>27.667</td>
<td>2.326</td>
<td></td>
</tr>
</tbody>
</table>

Table P. 2  Means and Standard Deviations of Overall Oral Performance Scores (P2) as a Function of Subtitling and Task Type

<table>
<thead>
<tr>
<th>Description</th>
<th>Subtitles</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>No-subtitles</th>
<th>Mean</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Subtitles</td>
<td>32.818</td>
<td>2.040</td>
<td>34.500</td>
<td>.874</td>
<td></td>
</tr>
<tr>
<td></td>
<td>No-subtitles</td>
<td>27.533</td>
<td>2.326</td>
<td>28.167</td>
<td>4.622</td>
<td></td>
</tr>
</tbody>
</table>

273

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Appendix Q

Means and Standard Deviations of Questions
("A Game"/"A Quiz") Scores (P1, P2)
as a Function of Subtitling and Task Type

### Table Q. 1
Means and Standard Deviations of Questions ("A Game"/"A Quiz") Scores (P1) as a Function of Subtitling and Task Type

<table>
<thead>
<tr>
<th></th>
<th>&quot;A Game&quot; (Description)</th>
<th>&quot;A Quiz&quot; (Narration)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Dev.</td>
</tr>
<tr>
<td><strong>Subtitles</strong></td>
<td>10.67</td>
<td>1.557</td>
</tr>
<tr>
<td><strong>No-subtitles</strong></td>
<td>9.33</td>
<td>3.114</td>
</tr>
</tbody>
</table>

### Table Q. 2
Means and Standard Deviations of Questions ("A Game"/"A Quiz") Scores (P2) as a Function of Subtitling and Task Type

<table>
<thead>
<tr>
<th></th>
<th>&quot;A Game&quot; (Description)</th>
<th>&quot;A Quiz&quot; (Narration)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Dev.</td>
</tr>
<tr>
<td><strong>Subtitles</strong></td>
<td>9.27</td>
<td>2.724</td>
</tr>
<tr>
<td><strong>No-subtitles</strong></td>
<td>6.80</td>
<td>2.484</td>
</tr>
</tbody>
</table>
VITA

A native of Valencia (Spain), Isabel Borrás is an elementary and secondary school teacher who has taught in both her country and France.

Isabel's education includes a B.A in Modern Philology from the University of Valencia (Spain, 1976), a M.A. in French from the University of Valencia (Spain, 1985), a D.E.A.V. (Diplôme d'Etudes Audio-Visuelles) from the University of Aix-Marseille (France, 1982), and a D.E.S.S. (Diplôme d'Etudes Supérieures Spécialisées en Technologie Educative) from the University of Poitiers (France, 1987).

In 1988 she came to the United States to pursue a doctoral degree in Education at Louisiana State University. In 1991-92 she designed and developed a multimedia package, "Practicing Spoken French," for her doctoral research project. While completing the requirements for her doctoral degree she has coauthored several publications and she has made presentations at both local and national professional meetings.
DOCTORAL EXAMINATION AND DISSERTATION REPORT

Candidate: Borras, Isabel

Major Field: Education

Title of Dissertation: Effects of Subtitled Video during Transactional Task Practice on Oral Communicative Performance of Fifth Semester College Students of French Learning with Multimedia Courseware

Date of Examination: January 26, 1993

Approved:

[Signature]
Major Professor and Chairman

Dean of the Graduate School

EXAMINING COMMITTEE:

[Signatures]

Date of Examination:

January 26, 1993