

# A New Challenge: Testing the Video Course

Malcolm J. Benson

Hiroshima Shudo University

Valerie A. Benson

Suzugamine Women's College,

Hiroshima, Japan

## Abstract

*This paper offers some preliminary thoughts on the testing of video courses, including a review of the literature and an examination of the research on the subject. The unique characteristics of video courses suggest that different language learning outcomes may be expected, which means that creating appropriate tests is therefore a new challenge for language teachers. The paper lists subtest types, and gives an example of a test used by the authors. Based on this experience, a number of guidelines are offered regarding the creation of appropriate tests, and a direction for future research in the area is suggested.*

## Introduction

The language learning possibilities of video courses<sup>1</sup> were first demonstrated to language teachers through a number of enthusiastic handbooks (e.g., Allan, 1985; Lonergan, 1984; McGovern, 1983; Stempleski & Tomalin, 1990) which tended to focus uncritically on their possible "applications." The pedagogic impact of video courses was rarely evaluated, though an early survey review (Bevan, 1986) did offer criteria for judging video packages. A more recent survey goes slightly further, suggesting that "attention should be focused on limitations as well as advantages" (Strange & Strange, 1991, p. 335). Articles now appearing show a further narrowing to specific uses: listening comprehension (Liskin-Gasparro & Veguez, 1990), productive language use (Cooper, Lavery & Rinvoluceri, 1991), literature (Bouman, 1991) and content-based teaching (Piñero, 1991).

However, there has been insufficient recognition at the theoretical level (except from Altman, 1990, 1989) that video courses have inherently new characteristics. The assumption,

*"Altman (1990) describes the present situation regarding video courses as 'the end of the beginning'...."*

presumably, is that video courses simply use alternative materials, and that some different classroom procedures may be necessary. Regarding testing, it is apparently felt that traditional testing procedures continue to be suitable.

The situation is complicated by the fact that video courses differ considerably from one another, and further, that even the same video material may be given quite different treatments in different pedagogical settings. Bevan (1986) distinguishes three types: course videos, resource videos, and supplementary videos. Strange and Strange (1991), perhaps reflecting the increasing sophistication of this market, offer five: course videos (also usable as supplements), "free-standing" videos which may or may not be related to other course materials, cultural videos (e.g., British life and culture), ESP (English for Special Purposes) videos (e.g., business English), and videos for children. Another complicating factor is the intensity of use: how often are the students given video time, compared with the time allotted for other kinds of instruction, explanation, exercises, and so forth?

A comprehensive assessment of all the above variables is quite outside the scope of this paper. However, it is hoped that the focus on the principles and practice of testing video courses will be sufficiently generalizable to offer something to all video course teachers. Regarding principles, we review and extend to language testing the ideas expressed recently in a number of papers; regarding practice, we offer a framework and examples based on ongoing experiences. This paper represents, essentially, preliminary thoughts and examples which others may find constructive.

## Recent Thinking About Video Courses

Altman (1990) describes the present situation regarding video courses as "the end of the beginning" (p. 9), meaning that we are now in the period following the initial excitement of a new teaching medium, but prior to its full integration into language teaching. Starting with schema theory ("understanding involves a dialectic between decoding and schematizing," p. 13), he moves on to argue that the intellectual shifts of the 1970s in literary and media studies have left a legacy in the discursive nature of all forms of television and video programs. All media productions nowadays are "targeted," and this aspect should be exploited by teachers: "Who is speaking? To whom? About what? To what end? What kinds of statement are people like this likely to make? What type of vocabulary do they use?" (p. 13). Here is Altman's central thesis:

Video provides a perfect vehicle for sensitizing students to discursive concerns, because it not only provides language, but also shows the circumstances of production of that language. With a literary or journalistic text, we have to understand the language in order to predict the language. Video ruptures this vicious circle by showing us a character's face before we hear her speak. Once we know that she is outraged by what she has heard, we easily predict the type of speech that she will utter. Because we know *to whom* she is speaking and *why* she is speaking, we more easily foresee *what* she is saying. (1990, p. 14)

According to Altman's thesis, video acts as an enhancer of understanding. However, Tuffs & Tudor (1990) question the "intuitive plausibility" (p. 30) of this type of thinking, arguing that non-native speakers are unable to avail themselves of the same range of contextual clues as native speakers. They suggest that the potential help from visual clues to meaning may be of "almost no benefit" (p. 43) in the case where the learner's culture is far removed from the target culture.

This debate symbolizes the ambiguous place of video courses in language-teaching curriculums. Weiland (1991) reports on a German project to assess the place of video in the language curriculums of secondary schools and universities. Video was found to stimulate discussion and group work, though note-taking and presenting results were unpopular activities. Nevertheless, the participating students achieved generally above-average results in English at the end of their course.

Berdahl & Willetts (1990) also outline some curricular concerns by asking questions such as: What kind of listening comprehension practice will a video course provide? Will it be used to introduce, expand, or reinforce a linguistic, sociolinguistic, or cultural concept? What objectives (e.g., grammar, functions, vocabulary, cultural situations) will be met? Teachers alone can address these theoretical points, and must then make practical decisions on classroom use, such as whether to use the entire video or just selected segments, whether to repeat certain segments, and so forth. The problems of "integrating" (the word is from their title) video into the classroom should not be minimized, and the authors finally recommend that "teachers need to take another look at

the total curriculum" (p. 22) before deciding on the role of video in it.

Within the curriculum, listening comprehension appears to be the skill that benefits most from video courses. Reporting a small-scale study, Terrell (1993) says that, "The data indicate that listening comprehension training with [authentic] video materials for intermediate students gives them experiences that cannot be duplicated in traditional classrooms limited to instructor/student interaction" (p. 22). Rubin (1990), working with Spanish video, concludes that "video can serve as a haven to enhance listening comprehension if it is selected so that it provides sufficient clues for information processing" (p. 315).

Secules, Herron and Tomasello (1992) report on two video research projects. In the first, which focused on listening comprehension, classes that used a video course (*French in Action*, Capretz, 1988) scored higher than control groups both in overall listening comprehension, and in the specific areas of main ideas, details, and inferences. However, they did not score higher on either reading comprehension or writing ability. In the second project, which focused on structures (vocabulary, idioms and grammar), they found no significant differences between the video and the control conditions. They concluded that (a) "students in the video-based curriculum clearly had better listening comprehension than did students in the traditional curriculum," and (b) that "only one significant difference was found between learning linguistic structures from video and learning these via oral drill and flashcards: vocabulary was learned better by oral drill than by video" (p. 486). These results lead the authors to surmise that in listening comprehension there are "other skills" such as "learning to use non-pedagogical contextual cues, understanding native accents, keeping up with native speaker speed, and recognizing a wide range of lexical meanings" (p. 487). Their comments, however, do no more than confirm that the testing of apparently isolated facets of language is subject to an inevitable "muddying" effect.

This "muddying" effect, in testing discussed as *construct validity*, emerges from a project that compared the effects of video versus audio tapes as media for the testing of listening comprehension (Gruba, 1993). Starting with the videotape of an academic lecture, he created a second, audio-only version. The versions were then administered to two test groups of 45 students, but t-test results were not significant at the .05 level, and test reliability (.45) was felt to be "disappointing." This leads Gruba to question the idea of using video as a test

instrument, in this case video in which contextual cues are limited. The study suggests that video in which extralinguistic cues and cultural context were minimal will yield insignificant differences in testing from the same material on audiotape.

Current thinking on video could be summarized as follows: (a) video-based instruction has particular strengths in the area of general understanding and comprehension, and in promoting oral activities such as discussion; (b) it needs to be carefully integrated into the curriculum, both on the theoretical and practical levels; (c) listening comprehension appears to be the greatest beneficiary of video-based teaching, but certain other areas (e.g., vocabulary) may show less progress; (d) no work has yet appeared on how to test a foreign language video course, though one project suggests that video may not be a suitable substitute for audiotape as a test of listening comprehension.

Benson (1993) has drawn together (from sources such as Altman, 1990; Herron & Hanley, 1992; Weiand, 1991, as well as from classroom experience) seven common claims made for all video courses:

1. The visual medium activates alternative learning styles.
2. Both decoding and schematizing are enhanced.
3. Video has a discursive nature.
4. Video allows access to nonverbal language.
5. Video lowers the affective filter.
6. Video allows access to the target culture.
7. Material seen on video is readily retained.

These claims have considerable implications for teaching with video courses, and are also fundamental to their testing.

## **Video Test Construction: A New Challenge**

If a particular video course addresses even some of the claims mentioned above, and also favors certain skills, as the research is beginning to show, then it follows that the testing of a video course presents a new challenge. Clearly it will be necessary not only to test what students have actually learned, but to test it in ways that have what Underhill (1987, 105) calls "face validity": "Does it look like a reasonable test? Do the people who use the test think it's a good test?"

In discussing recent trends in language testing, Byrnes (1992) draws attention to the importance of audience, of context, and of a range of sociolinguistic factors which must be taken into consideration. She uses the term "prochievement testing," which she offers as a combination of "proficiency" and "achievement." It can, she says, "capture the interactive, purposeful, audience-driven, creative use of language which

is the hallmark of natural language use" (p. 23). Prochievement testing can work by "identifying limited tasks which can challenge our learners' ability to use the language in a valid context but which do not presuppose total command of the language" (p. 23). This criterion-referenced approach to test creation was selected here as it speaks directly to teachers working with video.

It is assumed, moreover, that teachers will follow some version of the traditional sequence of looking at broad goals, creating teaching objectives to fit these goals, teaching to these objectives, and finally testing what has been taught. Also to be considered are the standard questions relating to the nature of any test: What kind of test is it (achievement, diagnostic, proficiency)? What competencies are to be evaluated (productive or receptive)? What decisions will be based on the results? And, given the variety of video courses, how do any or all of these questions differ across the various types of courses currently available?

## **Test Methods**

A large number of subtests is available, designed to test (in Byrnes' word) the students' "prochievement" in specific areas such as oral skills, comprehension, listening, grammar, writing, culture, and so forth. These have been broadly classified into productive and receptive skills. Below is a partial list (from Benson, 1993).

### *Productive Skills Using Video as a Stimulus*

- Retelling the story
- Oral production based on a frame from the video
- Repetition of sentence/dialogue
- Reading the dialogue (in pairs, etc.)
- Roleplay based on the dialogue
- Reading aloud from a parallel script

### *Receptive Skills Using Video as a Stimulus*

- Summary with clozed blanks
- Describe (in writing) what happened
- Arrange pictures (stills from the video) in order
- What does the conversation refer to?
- What is the main topic of the conversation?
- Continue the story from one character's viewpoint
- Matching exercises of various kinds
- Forms, etc., to be completed from the video
- Brief answers to general comprehension questions
- Multiple choice items
- Transcoding information from video to another medium

- Re-presenting the material in writing
- What colloquialisms/idioms were used?
- (Sound off) Express same ideas in different words
- Translation of the script into student's L1
- (Sound off) What kind of conversation are they having?
- (Sound off) What is A's attitude towards B?
- Reporting speech (Reported Speech lesson)
- Sentence transformation
- Write a short description of a person or place
- What "unusual" cultural aspects did you notice?
- Was the behavior polite or rude?
- Identify locations
- Contrast specific cultural aspects
- Match words + pictures
- Who ordered what from a menu?
- Identify the speaker—who said it?
- What specific facts were mentioned?
- What verbs were used?
- What word was stressed?
- Sentence correction: "That's not what was said!"
- Complete a transcription (clozed blanks)
- T/F questions—either details or main ideas
- Complete the other half of the dialogue

Now, it can be seen that *not* all of these are either "new" or "challenging," or even in direct accord with the principles outlined in the previous section. Some, for example multiple choice items, are totally traditional. Nevertheless, from this list a principled selection could be made, appropriate to an individual class. Let us see how this works in practice.

### Example of a Video Test

The following description concerns the testing of a year-long video course at a Women's Junior College English Program in Japan. The students were female second-year English majors, with an average age of 19. They had completed six years of formal English instruction at school, followed by one more year of Junior College English. Their general level would best be described as intermediate. The video course described here was one of six English classes that they were taking in their final year.

The broad institutional goals of the course were to increase the students' communicative competence in English. In practice this meant an uneven distribution of the four skills, with listening and speaking taking precedence over reading and writing. Students were also expected to acquire elements of the English-speaking culture. To achieve these goals, the right

combination of motivating material with suitable tasks at an appropriate structural and lexical level, and with sufficient cultural exposure, had to be found.

The choice fell to the teaching video *Mystery Tour* (Viney & Viney, 1988), which is best described as a “free-standing” course based on a simple thriller plot, set in Oxford, England. As the action unfolds, various teaching points (e.g., giving directions) emerge from the script. The accompanying “activity book” focuses the student’s attention on listening skills, grammar points, functions, expressions and vocabulary, as well as giving them opportunities to do roleplays based on the script. The class met once a week for 90 minutes, the Japanese norm. The students viewed the video material during class, but did not have access to it outside of class.

Next, course objectives were created for the receptive areas of listening, reading, and cultural observation, and for the productive areas of speaking and writing. Their construction followed Nunan (1988, p. 65), who has suggested that objectives should be described in three elements: tasks, conditions, and standards. (See Figure 1.)

---

<b>Specific Objective:</b>	When listening to native-speaker exchanges students will be able to identify the main topics under discussion.
<b>Task:</b>	Identify the main topics discussed.
<b>Conditions:</b>	While watching the exchange on video.
<b>Standards:</b>	Seventy-five percent of the main points should be correctly identified.

---

Figure 1. A *Mystery Tour* Listening Objective.

Despite their forbidding appearance, objectives provide considerable freedom, and attainable targets for all the students. They also greatly facilitate both teaching and testing (Nunan, 1988, Ch. 5). In this case, the course objectives (see example at Figure 1) led directly to the construction of the following test:

***Productive Skills based on the Video***

1. Roleplay based on the dialogue (everybody).
2. Retelling the story (half the class).
3. Description based on frames from the video (half class).

***Receptive Skills (Comprehension of the video)***

1. Listening: “Was it mentioned?” “Were these words used?”



2. Writing: Write a description of a character just seen.
3. Grammar: Given answers, what were the questions asked?
4. Culture: Judge polite/rude behavior; contrastive cultural aspects.
5. Comprehension: Rational cloze on content of first unit.

### The Productive Test (Oral)

The oral component of the test allowed the students (a) in the Roleplay, to demonstrate the extent to which they could handle certain structures seen in the video and practiced in the associated classwork, and (b) in the Retelling the Story or Describing, to show their ability to produce language based on what they had seen.

The practical sequence was as follows: The students were asked to sign up in pairs to be tested. In the testing room half the students were given Description plus Roleplay, and the other half Tell the Story plus Roleplay. Roleplay, being more demanding for Japanese students, was always done second, after the students had warmed up. In the Description task, the students were shown selected stills from the video which are reproduced in the activity book, and were then asked to describe what they saw (Fig. 2).

---

Teacher: (showing still picture from workbook) A, we're almost at the end on Episode Eight-, Episode Eight.

Student A: David

T: This is David, right. What is he wearing there, what does he look like?

A: He's wearing the jacket (pause), grey jacket, and white shirt and red tie.

T: What about Ros?

A: Ros is wearing a black, black shirt and a dark green jacket, and she has fair wavy hair.

T: What- Who was she talking to?

A: She is talking to, to David.

T: What was she telling him?

A: She is telling him that- to do what she said.

T: Right...(continues)

---

Figure 2. Example of Description.

Description was scored by one of the writers on nine criteria: Clothes, People, Appropriate Structures, Size of Utterance, Independence, Hesitation, Complexity, Speed, and Pronunciation. A nine-point scale was used, with three bands: low (scores 1 – 3), medium (4 – 6), and high (7 – 9). With nine as

the maximum per criterion, a total score of 81 was possible. The test results for the 25 students who did Description ranged from 18 - 69, with a mean of 42.84 (S.D. = 12.66). The example shown above scored 65.

The subsection Tell the Story was done by the other half of the class. The students were prompted by another still and were asked to tell the story leading up to, or forward from, the scene depicted. Figure 3 gives a short excerpt.

---

Student: Number one, there are Oxford [T.: You got it right.] and, Street of Oxford Tour, and tourist is very (inaudible) bus in front of the hotel, and so after that they are going onto the bus, the coach bus, and Nina was explaining about tour. When she was explained about tour, there are men whose name is Mr. Curtis, and he was late and so Nina noticed him and he's walking round and he's getting onto the bus. And she- in the bus Nina-, Nina interviewed some people and she also interviewed Mr. Curtis but he didn't reply *anything* at all. And so when she was found out why, and she was, she was telling about Mr. Curtis to Patrick who is the bus driver, and then she get-, then they got to the (pause) pl-, place (giggle), and Nina was explaining about that to the people, but Mr. Curtis is not interested in *anything* at all about tour.

---

Figure 3. Example of Tell the Story.

The 22 students who did Tell the Story were judged on six criteria: Basic Plot Told, Characters/Roles, Structures, Speed, Size of Utterance, and Pronunciation. At nine points per criterion (as above), the possible total was 54. In the test the marks ranged from 18 to 54, with a mean of 40.73 (S.D. = 7.66). The account shown above scored 47.

All students did Roleplay. Roleplay had both grammatical and functional constituents, being based on specific structures and functions that had been taught. The function chosen for this test was "suggestions," as it had been dealt with extensively in class. The students were given paired cards containing guidelines for the conversation, for example, Card A: "Okinawa/2 nights, 3 days/Sightseeing/March (after graduation)," and Card B: "You want to go, but how much will it cost? You might not be able to." This produced dialogues like the the in Figure 4.

Five criteria were used to judge Roleplay: Suggestion conveyed/understood, structures, speed, register, and pronunciation. The maximum on this subsection was therefore 45. In

the example above, student "A" above scored 36, and student "B" got 21. Test results were satisfactory ( $n = 47$ ; mean = 26.28; S.D. = 8.02).

- A: B, how about going to Okinawa for two nights and three days, to sightseeing, um, I think, I think March is, it's convenient for us to go in March, after we graduate (giggle). How do you (unclear: [think/feel])?
- B: I want to go to Okinawa in March, but...how much will it cost?
- A: Cost, (long pause) maybe it cost about...five, about fifty thousand yen.
- B: I'm sorry, I, I might not be able to go...
- A: (long pause) Why?
- B: (long pause) The cost is, the cost is very expensive for me.

Figure 4. Example of Roleplay Dialogue

### The Productive (Oral) Tests: Comments

Description was done relatively well. Students tackled it with interest and enthusiasm, scoring well on descriptions of clothes, facial features, and hair. Tell the Story, also enthusiastically done, produced the highest mean score and therefore might not be so useful if discrimination between the students were important (which it wasn't here). Roleplay was difficult for these students, and they did not perform well, often barely rephrasing what was given on the prompt card (see Fig. 4). In mitigation it could be said that the genre of roleplay is, of course, a culturally unfamiliar one to Japanese students, and good performances are rare.

### The Receptive Tests

Five written subtests were created: listening, writing, grammar, culture, and comprehension. For each of these a large number of choices was possible (see Section 3.1) and choices were made not only in line with the principles outlined in Section 2, but also with an eye to practicalities. For example, listening was offered in the "Was it mentioned or not?" and "Who said it?" format, as shown in Figure 5.

On the theoretical level we can relate this type of test to the ideas on decoding and schematizing (Altman, 1990) discussed earlier; on the practical level it is a direct test of an objective set out at the beginning of the course. The results of this ten-item subtest (one point per item equals a possible 10 total) on the group ( $n = 52$ ) showed the comparatively high mean of 7.8 (S.D. = 1.7), which was close to the notional criterion level of 75% suggested in Figure 1 above.

---

You are now going to see Episode 2, Section 2. Watch it carefully, and then answer the following questions.

1. Was it mentioned or not?
 

(a) A fire in a clothes shop.	Y N
(b) David went to the shop himself.	Y N
(c) David had a very successful morning.	Y N
(d) Nina remembers learning to make i's at school.	Y N
(e) David gave information to the police.	Y N
  
  2. Were these words used by these people?
 

(a) Nina: Did you say "bubbles"?	Y N
(b) David: Three cans of petrol?	Y N
(c) Nina: You are silly.	Y N
(d) Nina: And so you told the police?	Y N
(e) David: The uniform was awful.	Y N
- 

**Figure 5. Example of Listening Test**

Writing was tested immediately after the video sequence used for listening. Students were instructed to "Write a short description of David as he appeared in the video sequence you have just seen. First, mention his height, size, and appearance. Then write about the clothes he was wearing." This task was directly related to a teaching objective concerning descriptions, which had been covered at some length in the class. The idea of testing writing using the stimulus of the video itself was in line with the theoretical claim (See section "Recent Thinking About Video Courses") that video lowers the affective filter.

The writing task also foregrounded the enabling skill *grammar of descriptions*. However, grammar itself had been so prominent in the course that it was decided to test it directly. This was done by a traditional format, though the questions were all derived from the script, thus providing the student with a double opportunity to get a right answer. The grammar subsection of ten items produced a mean of 7.5, again just at the criterion level.

Cultural knowledge was tested by items which asked the students to make judgments about (a) whether something was polite or rude (e.g., a man in a pub in England calling out his order to the barmaid), and (b) whether such things could take place in Japan. This test, crude as it was, aimed primarily to test the students' cultural observation, but with

the secondary intention of having a positive washback effect. In the event, this subsection proved to be too easy, producing a mean of 4.25 (S.D. = 0.8) out of a possible five.

To test overall retention of the story, a rationally deleted cloze passage describing the first two episodes was used. The clozed words all expressed key ideas of the story, such as identities ("tour guide," and "boyfriend"), significant items (a blazer), and locations (Oxford's colleges). This ten-item subtest produced a mean of 7.0 (S.D. = 1.5), suggesting that the general structure (plot and characters) of the video had been well understood.

### The Receptive Tests: Comments

Every effort was made to test the students not just on *what* was taught in the course (the content), but also on the *way* it was taught (the medium). For our subtests of listening, writing, and culture this was easy in practice, and was also in line with Byrnes' (1992) idea that classroom testing should only be "an extension of what had already taken place in class" (p. 28). The use of the video as a stimulus for these subtests was received without comment by the students; we took this to be an indication of how naturally it followed on from the teaching. However, we did not feel we had found a satisfactory way to test grammar and comprehension, other than by traditional means.

### Conclusions and Future Research

Research on video courses is beginning to show a different profile from traditional courses regarding the learning outcomes that may be expected. Such courses now appear to offer benefits which, broadly, are clustered around the comprehension-listening skills, with less impact being felt in the traditional skills of reading, writing, and vocabulary development. Video courses have discernible characteristics which should lead teachers to select video only when the class goals are appropriate. At the very least, teachers need to recognize the different nature of the learning experience that their students will undergo.

*"Because language teaching curricula have been slow to absorb video courses into 'mainstream' teaching, video testing has also suffered neglect."*

Because language teaching curricula have been slow to absorb video courses into "mainstream" teaching, video testing has also suffered neglect. However, in this paper we have listed both productive and receptive subtests (see Section 3.1), and have suggested that a principled selection from this list would result in appropriate tests. Such appropriate tests are urgently needed because at present the learning outcomes of video courses are unclear, rendering tests a matter of hit and miss. Such testing, in turn, severely damages the credibility of video courses, preventing them from entering mainstream

teaching. This vicious circle can only be broken if teachers create valid and reliable tests for their classes.

The starting point for the creation of a satisfactory test is that the *purpose of the test* must be clear in the teacher's mind. Next, any test created for a video course should be *based on the course objectives*, and should reflect the teaching situation as closely as possible. The concepts "receptive" and "productive," while not always helpful in thinking about the development of communicative language skills, may well be useful in the testing of those skills. In addition, *the cultural side of the video should be both exploited and tested*.

A video course is more varied than textbook material, with the result that *different test-construction techniques must be employed* for each one. We have given a brief account of one attempt at test creation and have outlined the mixed results achieved. Unfortunately, no discernible pattern is evident at this point, and further research should consider the effects of different mixes drawn from the list of subtest types (listed on page 16 – 17). Teachers of video courses urgently need to establish some kind of interim testing model which can serve until a comprehensive paradigm of a video test is finally established. ■

The authors thank Paul Gruba of Kanda University of International Studies, Chiba, Japan, for his detailed comments on an earlier version of this paper.

- Notes**
- <sup>1</sup> The term "video course" is used throughout this paper as a shorthand for any professionally produced language-teaching course that has video as its main source of input. Other descriptors such as "video-based course," "video materials," etc., exist but are not used here. A video course of the type discussed here (*Mystery Tour*, Viney & Viney, 1988) is acted much like a movie, but with language that has been carefully scripted to place pedagogical interests above authenticity. For example, narration and description may occur frequently because of their value to the targeted students. The locations tend to be largely informal and everyday (taking place in offices, houses, etc.), but with a fast-moving plot to maintain student attention. The result is a pedagogical movie backed up by activity books, textbooks, and other supporting materials.

## References

- Allan, M. (1985). *Teaching English with video*. London: Longman.
- Altman, R. (1990). Toward a new video pedagogy: The role of schema theory and discourse analysis. *IALL Journal*, 23(1), 9 – 16.
- Altman, R. (1989). *The Video Connection: Integrating Video into Language Teaching*. Boston: Houghton Mifflin.
- Benson, V. A. (1993). Testing made-for-ELT videos. *The Language Teacher*, 17(9), 15 – 19.
- Bevan, V. (1986). Survey Review: Purpose-made videos for English language teaching. *ELT Journal*, 156 – 168.
- Berdahl, I., & Willetts, K. (1990). Steps in integrating video into the secondary foreign language curriculum. *IALL Journal*, 23, 17 – 24.
- Bouman, L. (1991). Video, an extra dimension in the study of literature. *English Teaching Forum*, 29(4), 15 – 17, 22.
- Byrnes, H. (1992). Recent trends in language testing: The case of testing oral language. *IALL Journal*, 25 (1), 15 – 29.
- Capretz, P. (1988). *French in action*. Hartford: Yale University Press.
- Cooper, R., Lavery, M., & Rinvoluceri, M. (1991). *Video*. Oxford: Oxford University Press.
- Gruba, P. (1993). A comparison of audio and video in language testing. *JALT Journal*, 85 – 88.
- Herron, C. A., & Hanley, J. (1992). Using video to introduce children to a foreign culture. *Foreign Language Annals*, 25(5), 419 – 426.
- Liskin-Gasparro, J. E., & Veguez, R. A. (1990). Teaching listening comprehension through video in first-year college Spanish. *IALL Journal*, 23(1), 37 – 49.
- Lonergan, J. (1984). *Video in language teaching*. Cambridge: Cambridge University Press.
- McGovern, J. (Ed.). (1983). *Video applications in English language teaching*. London: The British Council.
- Nunan, D. (1988). *The learner-centred curriculum*. Cambridge: Cambridge University Press.
- Piñero, C. H. (1991). Developing video materials for content courses. *Cross Currents*, 18, 196 – 200.
- Rubin, J. (1990). Improving foreign language listening comprehension. *Georgetown University Round Table on Languages and Linguistics 1990*, (309 – 316). Washington: Georgetown University Press.
- Secules, T., Herron, C., & Tomasello, M. (1992). The effect of video context on foreign language learning. *Modern Language Journal*, 76, 480 – 490.

- Stempleski, S., & Tomalin, B. (1990). *Video in action*. Englewood Cliffs, NJ: Prentice-Hall.
- Strange, J., & Strange, D. (1991). Survey: Video materials. *ELT Journal*, 45, 335 – 359.
- Terrell, T. D. (1993). Comprehensible input for intermediate foreign language students via video. *IALL Journal*, 26(2), 17 – 23
- Tuffs, R., & Tudor, I. (1990). What the eye doesn't see: Crosscultural problems in the comprehension of video material. *RELC Journal*, 21(2), 29 – 44.
- Underhill, N. (1987). *Testing spoken language*. Cambridge: Cambridge University Press.
- Viney, P., & Viney, K. (1988). *Mystery tour*. Oxford: Oxford University Press.
- Weiland, H. J. (1991). 'Schulfernsehen Englisch' - Sprachlernen mit dem Videorekorder (Projektbericht). [Television English in school - language learning with the video-recorder. Project report.] *Die Neueren Sprachen*, 90, 292 – 307.
- 

*Malcolm Benson is a Professor at Hiroshima Shudo University, Japan. He is currently co-editor of the JALT Journal.*

*Valerie Benson is a Professor at Suzugamine Women's College, Hiroshima, Japan.*