The role of visual media in listening assessment

Paul Gruba

Abstract

With a growing and widespread usage of video, satellite broadcasts and multimedia applications in the language classroom, test developers are increasingly challenged to provide for visual media in the assessment of second language proficiency. In tests of listening comprehension, particularly, the implications of the use of visual media as a presentation medium have yet to be fully explored. In brief, the paper discusses how the use of visual media as a mode of presentation might influence the definition, purpose and validity of second language listening comprehension instruments.

1. Introduction

In recent years, a number of studies have investigated multimedia usage in foreign language classrooms (Hunt, 1993). Multimedia applications, due to their ease of use, recognition of different paths of knowledge acquisition, ability to be student self-paced, thematic presentations and increased availability, are sure to hold a central position in the future of language instruction. None the less, the use of multimedia in the language classroom dramatically challenges the language educator in a number of ways. Primarily, the language educator is forced to reconsider what is to be taught. Kramsch, in a section entitled 'The challenge of multimedia' (1993: 196-204), writes

"The enormous educational potential of the computer is confronting teachers with their pedagogic responsibilities as never before. Never before have teachers so urgently needed to know what knowledge they want to transmit and for what purpose, to decide what are the more and the less important aspects of that knowledge, and to commit themselves to an educational vision they believe in." (1993: 201)

Multimedia, of course, is not just a compilation of several types of media, but a unique medium in its own right. The blending and juxtaposition of music, authentic spoken discourse, printed text,
photographs, video clips and graphics within a contextual environment make it a unique mode of presentation. For the present discussion, however, it is useful to view multimedia as part of a cluster of 'visual media'; that is, one of several media in which visual elements are primary attributes. Other visual media, for example, include video, television, video disks, variations of CD-ROM technology and teleconferencing. 'Visual media' is a preferred umbrella phrase due to the rapid expansion of technologies in this area; further, it allows the researcher to draw on a larger body of literature in the field and provides a category expandable for future developments. Importantly, the term 'visual media' follows work by Salomon (1979) on the 'non-notational symbol systems' of television. Such a conception of visual media allows it to be conceptualised in terms of its component elements, not as a holistic entity. From the point of view of the language researcher, it provides a way to see learners' interactions with visual media as a search for appropriate meanings across, and within, shifting contexts.

One way to delimit the role of visual media is to move the site for research from pedagogical applications in the language classroom to that of the language test: classrooms are such complex environments (Chaudron, 1988) that investigations set to determine specific influences of a medium are frustrated by the inability to control intervening variables (Salomon & Clark, 1977). The framework of language testing research here provides a 'set of tools' on which to clarify component parts and eventually build construct definitions (Bachman, 1990a). Fortunately, the strong interdependence of language teaching and language testing (Davies, 1990; Heaton, 1988) point to an applicability of the findings in one area to another.

2. Sound and image: Listening comprehension

In addition to an established use of audio-tape and printed words, language testers have employed a variety of media to present language, including video tapes (Baltova, 1994; Benson & Benson, 1994; Berry, 1995; Gruba, 1994; Pelletier, 1990; Ryberg, 1995), video disks (Rock & Cummings, 1994) satellite teleconferencing applications (Clark & Hooshmand, 1992; Oxford, Ito, & Sumrall, 1993), and a variety of computer applications (Dunkel, 1991a).
These visual media have, for the most part, been used to assess listening comprehension skills.

The skill of listening is poorly understood as research in listening has lagged behind other skill areas (Rubin, 1994). Given the large number of variables which may affect comprehension—attention, auditory perception, memory, the evaluation of messages, difficulties of texts, etc—listening can be seen as the "consummate multivariate topic" (McDevitt, Sheehan, Cooney, Smith & Walker, 1994: 232.) In light of new media, the use by researchers such as Berne (1992) of Wolvin and Croakley’s definition of listening as the "process of receiving, attending to, and assigning meaning to aural stimuli" (1985: 74) appears limited. Kellerman (1990, 1992) argues that the majority of second language listening comprehension taxonomies (eg. Lund, 1990; Richards, 1983) ignore the importance of visual cues in the understanding of speech and, as a result, are flawed. Indeed, others (Altman, 1989; Gassin, 1992; Hurley, 1992; Oxford, 1993; Pennycook, 1985; Rubin, 1995; Willis, 1983) put forward arguments that non-verbal elements are crucial to communication.

In this paper, listening comprehension is defined as the ability to receive, attend to and assign meaning to both aural and nonverbal stimulus in context. Such a view is supported by Ur who states that successful listeners a) listen for a purpose and have expectations, b) respond to what they hear, c) see the person to whom they are listening, d) see visual and environmental cues, e) process heard discourse in small chunks and, further, that f) most heard discourse is spontaneous and contains redundancy, 'noise' and colloquial expressions (1984: 9). Apart from specialised telephone conversation training, learners, Ur (1984) states, should be encouraged to rely on their ability to employ visual cues when seeking meaning (as when using their own language) and that teaching listening skills in visual isolation lacks value.

Rost (1990) places nonverbal signals under ‘editing strategies’ that are used to more successfully infer speaker intentions. For Rost (1990), it is the interaction of verbal, vocal and visual characteristics of a speech event that shape the perception of speaker intention. Rost explains that visual and gestural clues cannot be correctly interpreted in isolation: they either converge or diverge from a given verbal signal. By themselves, they only
provide "weak inferences" in relation to a speaker's intentions (Rost 1990: 79).

According to Dunkel, the study of listening is "a polestar of second language acquisition theory building, research and pedagogy" (1991b: 432) and is critical to a society entering a 'post-literate' era in which visual media diminish reliance on printed media and enhance the use of spoken texts. Dunkel (1991b) sees an urgent need for a model of communication that includes nonverbal signals, echoing Canale and Swain (1980) who write:

"More research on the role of such nonverbal elements of communication as gestures and facial expressions in second language communication may reveal that these are important aspects of communication that should be accorded more prominence in the theory we have adopted." (1980: 36)

One tentative framework specifically intended for use in listening comprehension assessment (Dunkel, Henning & Chaudron, 1993) sets the place for visual media under the rubric of 'levelling variables' (factors most likely to affect an examinee's score) as 'video support' with the implication that images serve to assist the understanding of aural messages. This view is, however, countered in studies by Baggett (1984) which see the image as the primary source of understanding, and the aural track attended to when an interlocutor has trouble with an unclear message.

3. Media in second language testing

Like other material developers, test designers use media to meet three practical concerns: control of the input, convenience and cost effectiveness. 'Live' presentations of a talk, for example, do not afford the language test designer sufficient precision needed to attain acceptable indices of reliability. Too many variables impinge on the ability to measure defined aspects of language competence for the instrument to capture a reliable score. Convenience, too, in that media are transportable, can be stored and archived, and can be further produced in multiple copies for international distribution. Cost control, always a concern in test production (Stansfield, 1993), is a particular benefit of media usage. It would appear that audio-tape is the clear winner in affordability. Purpose-built video productions are much more
expensive to make than audiotape productions (Ariew, 1987); multimedia production costs soar well beyond the budgets of most test developers.

Tests, however, unlike text books or other educational materials, have a relatively brief shelf life. Concerns that examinees may have seen the material prior to sitting for an examination and thus would potentially achieve a rating above a 'true score', point to a systemic bias in the instrument that would render it unacceptable for repeated use. As such, language test professionals are almost constantly embroiled in a cycle of development and redevelopment of their products: an expensive process.

So why use visual media in listening assessment? Critics could argue that audio tapes, less costly to produce, easier to administer and supported by a body of work in the area, work well for the majority of test purposes. Purpose, the first consideration in test development (Bachman, 1990b; Dunkel et al, 1993), points to the role of visual media in language assessment, in that the 'role' is determined by what the test developer would like to achieve with the instrument. There are four categories of purpose which visual media may address, none of which is mutually exclusive. To a large extent, the purpose categories proposed here are inter-related.

3.1 Theory driven

A belief that assessment instruments must better correspond to models of language comprehension that explicitly incorporate aspects of culture (Phillips, 1995), nonverbal elements (Gassm, 1992; Knapp, 1980), or other aspects of non-verbal communication including lip reading (Kellerman, 1990) drives the test developer to look beyond audio tape as a mode of presentation. Validity concerns (given the specialised nature of audio only interactions (Ur, 1984)) and the fact that audio alone "may be sufficient for those knowledgeable of a domain, but visual symbol systems supply important situational information for those less knowledgable" (Kozma, 1991: 192) raise problems related to background knowledge and language proficiency.
3.2 Pedagogy related

Films, created by Disney studios, were first used in foreign language teaching in the 1930s; television has been used in language teaching since 1947 (Kelly, 1969). Another reason to use visual media in assessment then is because we would like to integrate our assessment procedures with existing pedagogic practices. Lynch and Davidson, for example, use the term ‘reverse washback’ (Lynch & Davidson, 1994: 728) to describe the influence of teaching practice on language test development. Given the widespread and principled use of video in teaching (Stempleski & Tomalin, 1990; Stempleski & Tomalin, 1995) this approach argues that media used in teaching should be the same as those in testing (Benson & Benson, 1994). At present, there is “almost total exclusion of television materials from the assessment process” (Kelley, 1993: 61).

3.3 Process specific

A third reason to use visual media in language assessment is because there are features of the process, or setting, of how the language is being used which can not be separated from its meaning. In a sense, this touches on the area of ‘English for Specific Purposes’ video usage (Kennedy, 1983). In Melbourne, for example, recent migrant students in English language programs are being trained in computer ergonomics to prevent repetitive strain injury. These students learn to sit at a safe distance from computer screens and how to sit properly for long periods of time. The use of visual media in the assessment instrument may help to identify students who have mastered proper work practices.

3.4 Mode of delivery

A fourth reason to use visual media in assessment is driven by their existing use in distance learning programs (Clark & Hooshmand, 1992; Oxford, Ito, & Sumrall, 1993, Warriner-Burke, 1990) In such situations, again, television may be a mode of instruction; language usage is inseparable from its presentation. Correspondence courses can follow a similar path, in that videotapes are included in learning packages; video conferencing through networked computers raises the possibility of live, interactive speaking and listening tests.
4. Images: Visual literacy

To make sense of a modern environment, the ‘reading’ of symbols and their meanings is a necessary skill. It is undisputed that images have the power to affect human behaviour, for example, as in the case of campaign advertising (Biocca, 1991; Postman, 1985). Though attempts to define the term ‘visual literacy’ have been largely unsatisfying (Seels, 1994), the ability to decode a visual statement in a particular context may become, according to Dondis “one of the fundamental measures of education in the last third of our century” (1973: 19). In addition, some learners are better able to interpret visual cues than others (Salomon, 1979), perceptual habits may be culturally determined (Seagall, Campbell & Hersokovits; 1966; Tuffs & Tudor, 1990), individuals may perform better in learning tasks using particular media over others (Carbo, Dunn & Dunn, 1986; Sless, 1981), and pre-conceptions of a given medium may affect a learner’s willingness to actively engage with its content (Cennamo, Savenye & Smith; 1991; Salomon, 1984). As with the lack of assessment criteria in nonverbal communication (Pillar, 1995), no established criteria exist to assess individual competence in visual literacy.

5. Textual features and production factors

When designing a video-mediated test of second language performance, Gruba (1993) was faced with a dilemma: what does the native speaker teacher look like? Eventually, Gruba (1993) based his choice of actor on the boundaries of test purpose and sociocultural context to address his problem. Conducting the study in an academic setting, Gruba (1993) established through records in the university personnel office that students would most likely be instructed by middle-aged, white males and thus, accordingly, chose to present a middle-aged, white male as a key figure in his video production. Production concerns in turn affected instructional design issues. Reeves (1992) writes that criteria for visual media productions should attend to 1) context for use; 2) individual differences; and 3) interactions between individual differences and aspects of user interface.

Appropriate interpretation of a text, with a clear understanding of its constituent elements that contribute to an overall theme or controlling idea, is a key step in language test design (Bachman,
1990b). What elements constitute the 'text' of a multimedia production? Further, how can a reasonable interpretation be made? Multimedia, in this regard, poses a challenge to the researcher because of its ability to blend, shift and juxtapose. Unlike the established genres of television (news broadcast, game show, soap opera), based on a linear mode of delivery (neither the broadcast, nor its rate of delivery, can be manipulated by the viewer and videotape is an awkward attempt to rectify this shortcoming), a multimedia application by its very nature changes structure each time it is accessed. The inherent instability of a such a medium poses a number of design issues (Yankelovich, Smith, Garrett, & Meyrowitz; 1988). In addition to appropriate content and task considerations (Nunan, 1989), the medium requires that designers attend to the pacing of text appropriate to the target ability group (Kozma, 1991), colour coding (Chapman, 1993), subtitling (Borras & Layfayette, 1994) and individual uses of the medium (Anderson-Inman, Horney, Chen & Levin, 1994).

6. Research: Methodological considerations

Gassin (1992) states that researchers interested in nonverbal forms of communication must, first and foremost, create a methodology to investigate such language forms. What would such a methodology entail?

First, it is important to distinguish between research on media and research with media; the latter preferred as a way to better inform educational practices (Salomon & Clark, 1977). One goal in research with multimedia would be to establish how second language learners go about using various media. Learner strategy researchers employ surveys, direct observation and think-aloud protocols as a means to establish differences in the ways individuals interact with language teaching styles and materials (Oxford, 1989; Oxford, Lavine & Crookall, 1989). An understanding of strategies learners display when faced with visual media would illuminate the cognitive processes that lie behind learner actions, perhaps yielding insights into differences in learner performances.

Strategy research has traditionally involved observation of learners in a classroom, but work by Cohen (1994) and suggestions by Kozma (1991) point to the successful use of introspective and restrospective protocols (Ericsson & Simon, 1984) as a means of
gaining valuable insight into the process of how learners go about learning with media. The real-time cognitive processing demands that learners face when working with multimedia applications on a second language may decreases the effectiveness of these protocols, though as Ericsson and Simon (1984) point out, the method is a viable procedure to investigate mental processes that may otherwise be lost. The shortcomings of this method force the well-rounded investigator to employ a number of complementary methodologies (Cohen, 1994) including participant observation (Parideen, 1991), text analysis (Fields, 1988) and semi-structured interviews (Buck, 1990).

7. Summary: Towards an agenda for research

Are tests based on visually mediated materials feasible? According to Spolsky (1995), three factors influence the development of a valid ‘post modern’ test: feasibility, useability and reliability. To this list, ethical questions should be added.

The question of test validity is perhaps best determined by the setting in which the instrument is to be used. At the University of Melbourne, there is access to fully-equipped lecture theatres which are capable of displaying not only video, but multimedia applications, to hundreds of examinees in a sitting. The lecture theatres are equipped with a range of speakers that attain professional sound quality. Additionally, the University has equipped a number of language classrooms with large screens and multiple television monitors in conjunction with several dozen high end computers. Technological feasibility, leaving aside other problems for the moment, is possible at a number of educational institutions.

A second factor in Spolsky’s (1995) checklist, the useability of visual media, is tied to setting and experience. In my development of a ninety minute video-mediated test in Japan (Gruba, 1994), the training of university administrators, instructors and students in use of the instrument went forward without significant problems.

The attainment of acceptable reliability indices, however, will cause delays in the development of valid assessment instruments that use visual media. A number of factors contribute to this delay: 1) listening tests pose a tremendous set of variables that make them
notoriously difficult to validate, and the inclusion of more variables further complicates this; 2) in a non-notational symbol system, there is room for multiple interpretations of a given event, further confronting the test developer with the need for appropriate textual interpretation; 3) research in learner strategies with regard to visual media use is sorely lacking, and there is no guarantee that candidates will attend to specific visual cues in a field of images; 4) researchers need to better define conceptualisations of validity in regard to 'culture' and 'context' in a post-structural environment. In short, the introduction of visual media to instrument design makes the complex task of assessing second language communicative proficiency further complicated.

A number of ethical issues arise when considering the use of complex technology in a number of educational settings (Cameron, Frazer, Harvey, Rampton & Richardson, 1993; Howe, 1994; Morgan, 1995). Visual media raise the possibility that the uglier aspects of racism and ethnic stereotypes will be compounded when displayed in educational settings (DeVaney, 1994; Ellsworth, 1989; Nicholas, 1994). However, it must be noted that neglect in the development of visually mediated instruments may, in fact, hinder a more widespread application of learning technologies (Dwyer, 1994). It is clear that post-industrial societies are moving towards further integration of computers in education: should the language tester strive to keep up with these trends?

This paper has attempted to discuss a number of key areas which are ripe for investigation in regards to visual media usage in foreign language assessment programs.

1. Visual media were presented as complex and unstable blendings of image, sound and production factors that combine to make them particularly challenging to the researcher. They are perhaps best understood as non-notational symbol systems. As a result, this paper suggested that a framework of language testing, as opposed to the classroom, be employed to delimit possible confounding variables.

2. Listening comprehension was reviewed in regards to nonverbal communication, and it was noted that language specialists have realised the importance of this aspect of successful communication but remain frustrated in their attempts to assess its impact. A
number of reasons why visual media may be needed to assess language were presented.

3. The relation of the test developer to media was discussed, and control, convenience and cost effectiveness were put forward to justify their current usage.

4. Visual literacy was offered as another possible avenue of exploration, but developments in this form of literacy are still at an early stage.

5. Methodological issues were raised and it was suggested that think aloud protocols, complemented by direct observation, extended interview and text analysis would assist investigation in this area.

6. Spolsky's (1995) framework regarding test validity was applied to visually mediated instruments.

8. References


