

LISTENING SKILLS DEVELOPMENT THROUGH MULTIMEDIA

Carla Meskill, University at Albany, State University of New York

Abstract

As multimedia technology (interactive videodisc, CD-ROM, CD-I, etc.) becomes more accessible to teachers and learners of other languages, its potential as a tool to enhance listening skills becomes a practical option. Multimedia allows integration of text, graphics, audio, and motion video in a range of combinations. The result is that learners can now interact with textual, aural, and visual media in a wide range of formats. Consequently, when we now look at the computer as potentially supporting listening skills acquisition, we need to examine not only *aural* processing opportunities, but *multi modal*, (simultaneous sight, sound, text) processing as well. This paper examines multi modal processing and its implications for listening skills development in a foreign or second language. How multi modal processing as it relates to listening skills development can be supported by multimedia technology is presented.

Background

The past two decades have brought to language teaching and learning a wide range of audio-visual technologies. From among these, no single tool for teaching and learning has had greater impact than the personal computer. Today, individual learners can, in addition to interacting with computer-generated text and graphics, control combinations of analog and digital sound and images. Arranging these combined media into intelligent, pedagogically-driven material is a challenge to materials developers. Effectively integrating the technology into language learning contexts represents a challenge for language teaching professionals. A critical step in accomplishing these goals is careful examination of the technology's features in light of the needs, goals, and processes of language learning. The following discussion is an attempt to focus attention on the multi modal features of the technology that can interact with the development of listening skills in a second or foreign language.

-Listening

In face to face interaction, listening entails complex interpretive processes. An intricate web of situational variables interact to determine what meanings are derived in conversation. Processing requirements such as reciprocity of interlocutors' perspectives, the etcetera principle (filling in the gaps of what one hears with knowledge of the language and the world), and combined retrospective and prospective meanings all come into play. This multi-faceted

processing spells a heavy demand when the medium of communication is a foreign or second language. Theoretical models that attempt to capture the intricate nature of the listening process cannot hope to account for the myriad of cognitive and external environmental factors that influence reception, interpretation, and response construction. In short, rendering a complex activity like listening into a single construct has proved difficult (Dunkle, 1986). Models that have been attempted, however, share one underlying assumption: Listening is not simply a receptive act -- multiple physiological and cognitive processes are engaged simultaneously.

Until recently, listening comprehension activity in foreign or second language classrooms was limited to *testing* listening comprehension. The underlying rationale was that if students are successfully learning the target language, they should automatically be able to decode the aural version of structures and vocabulary they learn in their textbooks. Success at this decoding was typically measured by correct responses to WH (information) questions. Responses to such questions tagged successful retrieval of information from an aural text. Knowledge of target language syntax and lexis was deemed sufficient to enable this retrieval and was ultimately how students were tested. Listening is now treated as a much more complex activity and one that is the cornerstone of language acquisition (Krashen, 1994).

Recognition of listening's critical role in the language acquisition process has greatly influenced contemporary language teaching practice. The view that listening as an active and interactive process has, for example, cast the learner in a role other than the passive receiver of aural input (Rost, 1993). Classroom emphasis is now on aural intake through active negotiation of meaning. In face to face interaction, the listener, not just the speaker, engages in the active making of meaning. It is believed that this mutual negotiation of meaning between speakers activates the cognitive and sociocognitive processes necessary for language acquisition to occur (Breen and Candlin, 1980; Pica and Doughty, 1985). In short, listening has been recast as an activity central to the L2 acquisition process (Dunkle, 1991; Krashen, 1981; Rost, 1993), and a skill integral to overall communicative competence (Brown, 1994b; Savignon, 1991).

-Listening and Technologies

A prominent artifact of older beliefs concerning the role of listening in language learning is the language laboratory. The rationale for language laboratories is tied to the belief that individual listening practice with audiotape can help build a learner's overall ability in the target language through self-instructed comprehension practice. Technology continues to be perceived as an enhancement to the process of language acquisition. The large-scale infusion of computers in language instruction programs in the past decade attests to this belief. The rationale behind what

is now growing support for Computer Assisted Language Learning (CALL) is not unlike earlier enthusiasm for audiotape-based technologies. That is, individualized access to target language material under learner control provides needed exposure to and practice in the target language. Enthusiasm for CALL in general and multimedia in particular, however, differs from that of the audiotape laboratory as regards the breadth of expectations concerning technology's role and potential. Fast and powerful computational capacity in conjunction with the orchestrated video, text and graphics of today's multimedia learning systems would predict more sophisticated paradigms for interaction with the target language and, consequently, more effective learning. Arguments supporting multimedia for education of this kind have rung loud and clear over the past decade. Praises for the medium are, however, based largely on intuition: learning a language via individualized instruction with the computer -- especially when audio and video are involved -- is an extremely appealing proposition, one that has sold to many an administrator in search of instructional panaceas. Thus far, however, the extent of multimedia's impact on the language acquisition process remains an open issue. Is there evidence to suggest that listening skills development can be enhanced through this medium? The following section treats this question by examining potential correspondance between multi modal processing opportunities for language learners and how these can interact to complement listening skills acquisition.

Multi Modal Processing

Multi modal processing refers to the engagement of more than one perceptual modality at a time. The opportunity for multimedia users to process combined media (text, sound, and video) simultaneously is a popular trend in software design in general, and language learning materials in particular. Proponents of instructional multimedia have vigorously argued that the increase of sensorial input available via the technology coupled with the potential for active engagement in, and interaction with this input predicts that content (in this case the target language) will be more readily integrated into a learner's developmental system and, in turn, recalled more thoroughly (Stevens, 1989; Underwood, 1990). Is the engagement of multiple modalities appropriate for language acquisition?

The question of whether and in what ways aural processing is enabled or impeded by additional and simultaneous forms of input has been indirectly treated in studies with native speakers and, in rare cases, with students of a foreign or second language. The prevailing arguments in support of and against multi modal processing as an aid to listening and second language learning are presented below.

- Visuals

Learning theorists have long held that images enhance comprehension, storage, and recall of information (Pavio, 1965). In the language classroom, use of visual material of all kinds has been a predominant tool for instruction for quite some time. Pictures, slides, drawings, and the like serve many roles in language learning activities. In listening skills development, activities that focus learner attention simultaneously on visuals and accompanying aural input are common. Visuals support comprehension and form-meaning correspondence, both of which contribute to higher levels of learner motivation. Where combining input modalities in the classroom is based largely on intuition, a handful of empirical investigations on the effects of combining perceptual modalities support the use of text and visuals as aids to aural skills development. Support for this sort of multi modal processing is comprised of evidence pertaining to the positive effects of visual accompaniments to the listening process. Evidence suggests that processing aural texts in the target language is facilitated by co-occurring still pictures (Mueller, 1980 Omaggio, 1979), video (Snyder & Colon, 1988) as well as combinations of visual, aural, and textual forms of input (Brownfield, 1990; Chiquito, 1994; Garza, 1991; Koskinen & Gambrell, 1993).

Including visuals for listening skills development also finds support when rates of spoken language and the human ability to process incoming aural information are considered. We process what we hear much more quickly than the time it takes for the message to be spoken. While we listen, we have time to infer and elaborate. When the language is our own first language, there is sufficient time and opportunity to mentally act upon the incoming stream by creating connections, making transformations, interpretations, and mental images. When aural input is in a language for which we have a limited ability, additional effort must be expended: a portion of the mental energy otherwise assigned to interpretation and elaboration gets channeled into challenging, unfamiliar, and mechanical linguistic issues. There is evidence that, due to these unique L2 processing and channeling demands, memory span is shorter than when dealing with native language input (Call, 1985). Because listening entails the construction of mental representations and interpretations, it makes sense to supply the L2 learner with stimuli that support and even extend this process. Visuals can provide just such support. The information contained in pictures can mean that less cognitive energy gets expended on linguistic decoding, energy that can be channeled to other critical processes -- predication and elaboration, for example -- of the input. In sum, aural processing can be viewed as supported and facilitated by visuals. Visual support provides the learner hooks on which to hang meaning and make sense of the aural stream.

-Text

There is increasing evidence that verbatim, co-occurring text with video can aid second

language comprehension (Garza, 1991; Markham, 1989; Price, 1983). Video subtitles can serve as advance organizers that support and scaffold meaning as it occurs through the aural channel (Lambert, 1986). In this way the presence of text can "diminish the decoding load placed upon the learner by the unrefined audio signal of authentic [speech and] materials" (Jung, 1990:208-209). The nature of verbatim subtitle text, moreover, is directly aligned with the goals and processes of Communicative Language Teaching; that is, subtitle text is *performance text*, not constructed, reflective text. What individuals say is what gets subtitled. Because subtitle text is what is spoken on the video screen, it more closely resembles oral communication, not writing. It represents, therefore, a rare opportunity for language learners to experience approximations of oral language in both aural and written form.

-Video

Where learner response to video as an instructional tool has been positive across disciplines, it has been particularly strong for language instruction. This is partially due to the positive attitudes toward the medium with which learners are predisposed. As regards language learning, strong receptivity may also be linked to the ease of aural processing that visual accompaniment implies. Video can fill in gaps in aural comprehension which at once lowers affect and empowers the language learner.

Video is widely considered more powerful, more salient, and more comprehensible than other media for second and foreign language students (Brinton & Gaskill, 1978 ; MacWilliam, 1986; Tudor, 1987; Vanderplank, 1990). In rare empirical studies, video-based instruction is consistently preferred over other language learning activities (Secules, Herron & Tomasello, 1992) as well as over audio-only instruction (Pederson, 1988). Multimedia systems with video under learner control are also preferred other instructional activities (Brooks et al, 1992; Brownfield, 1990). In short, multiplying input modalities to include full motion video apparently motivates learners and engages their attention to aural input.

The co-occurrence of video with text, audio, and graphics in the multimedia environment raises issues concerning the amount of processing these input modalities entail and whether these demands limit or lengthen task persistence. First, our response to any medium is heavily mitigated by the extent of our experience with it. In the case of video, the role the medium has come to play in the lives of all contemporary peoples is extensive. Language students come to the learning process well versed in the medium and its conventions. They come literate and psychologically prepared to attend to and react to video using skills and strategies for understanding that they have developed over their lifetime. As a consequence of extensive prior

experience, users are motivated by the medium as well as accustomed to decoding its messages for extended periods of time. Second, the cognitive requirements of multi modal processing may also imply increased task persistence. As discussed earlier, there is a lag between the pace of aural input and the time required to process it. When one's first language is the medium of communication, mental elaborations take up this lag time; when the medium is a second language, demands for simple decoding are strong and happen at the expense of elaboration. However, if other forms of support for aural input are made available, attending to overall meaning derivable from multiple representations of input may take precedence over a preoccupation with form; e.g., comprehending individual words and sentences. Multi modal materials, then, potentially support comprehension of the message as opposed to drawing attention to its constituent parts.

-Schema

One aspect of language processing widely held as supporting and enhancing comprehension is that of mental schemata. Research in reading in both the first and second language support the notion that activating knowledge of the world and applying this knowledge to new input greatly facilitates processing and understanding. Good readers, for example, call on their past experiences and knowledge of the world when making sense of text. Likewise, when we process aural and visual input, our existing knowledge structures interact with incoming information (Luke, 1985; Salomon & Leight, 1984). Listening, like reading, is an active process that entails construction of meaning beyond simple decoding. Activation of what is known about the world clearly assists processing the aural code. The kind of processing considered optimal for language acquisition is bound to the complex contexts within which the aural text resides. Learners make use of this context through a process of matching new input with meaning based on their previous experiences (Diller, 1981). Learning to trigger and utilize mental schemata is an important strategy for language students to learn and employ.

In addition to drawing on direct experience and existing knowledge about the world, with multimedia learners can also be guided to capitalize on schema related to their media literacy; they can be prompted to effectively utilize their highly developed familiarity with visual conventions (camera angles, special effects, for example) to understand the wider context and, in turn, the aural text. Video in particular can set up a "context of expectations" that, like knowledge of text convention, can support comprehension (Salomon & Leight, 1984). Multimedia is an excellent medium for exploiting this feature of aural processing. Clearly, contemporary learners come to instructional experiences possessing skills and strategies for decoding and comprehending film and video. They may not, however, consciously realize that these

conventions can be cues to meaning associated with the aural text. In a multimedia environment, learners can be easily prompted to make use of such visual conventions to aid their aural processing.

Facilitating the activation of prior knowledge and the linking of old and new information can be achieved through any one, or any combination of processing channels: text, audio, visual. Attempts to assess the effectiveness of schema activation in one or a combination of modalities has supported the rationale for schema-related tools. Drawing learners to a context of expectations through various media combinations has been shown to assist comprehension and retention of aural input (Chiquito, 1994; Gay, 1986; Meskill, 1991c; Borrás, 1993). Additionally, one study of second language comprehension strategies undertaken when reversed subtitling was the medium demonstrates the scaffolding effects of mental schema in aural input processing. Lambert had subjects listen to input in their native language while reading the same text in the target language. Information supplied through the native language appeared to support the decoding and comprehension process of the target language text. The study concludes that the provision of information through a readily accessible channel (aural native language) serves to activate sufficient contextual information (mental/conceptual schema) for learners to more readily decode and comprehend second language input (Lambert, 1986). Support for enacting narrative schemata from studies in reading (Collins, 1981) and media (Baggett, 1979) also contribute to the notion that drawing learner attention to schema strategies they might employ while working with multiple forms of input may be beneficial.

Another aspect of mental schemata of particular interest to second language acquisition is that of scripts. Scripts are conversational *templates* or specific verbal routines that exist in all languages to accomplish certain communicative goals. Familiarity with scripted routines is quite automatic in one's native language, but potentially problematic for the second language learner. This "script competence" is a critical feature of successful second language learning, especially for successful comprehension of aural input (Dunkle, 1986; McCarthy, 1991). Developing familiarity with target language scripts - understanding how things get accomplished with the language via formulaic routines -- can also be realized within a multimedia environment. Students can be encouraged to uncover underlying conversational structures through combined media. Learners can, for example, be provided tools to access and manipulate text or pictorial representations of those scripts that underlie the aural text on the computer screen.

-Chunking

Where the basic unit of foreign and second language study was once discrete elements -

words and expressions isolated from a context - it is now lengthier stretches of discourse. The rationale for this shift in curricular focus is consistent with that of Communicative Language Teaching; the *context* in which words and structures get uttered plays a major determining role in relaying meanings for those words and structures. One cannot, for example, fully understand the intended meaning behind the present perfect tense nor the meaning of the word "plug" in isolation. It is the context in which these words get used and for what ends that render them substantive communication. Multimedia represents a unique environment in this respect. Aural text can be accessed, presented, and repeated according to predetermined breaks while the context (represented visually and even textually) remains immediate and at hand. Also, in considering multi modal processing, the learner's simultaneous processing benefits from such apportioning of the aural text. Language processing time can be allotted during breaks in the aural stream. In conjunction with the visual and textual, chunking renders the aural stream more manageable, with the needed processing time between chunks lengthened to what the individual learner deems necessary.

There are many approaches to portioning discourse for the purpose of listening skills development. One of these approaches is to chunk aural texts by syntactic breaks. These are the "breathing points" in utterances that are governed by the syntax of the language. When an aural text is chunked by these natural pauses, learners can be cued to use not only structural cues but the equally rule-governed cadence of the chunks as well; e.g., patterns of stress and intonation. In a study of second language learners' aural comprehension, O'Malley et al found that, unlike weaker listeners, effective listeners actively used intonation contours and syntactic breaks to chunk and process the aural text (O'Malley et al, 1989). Like with reading, good listeners make good use of chunks to understand what they hear (Hawkins et al, 1991). In a study with multimedia materials for French that were syntactically chunked, Chevillard's subjects reported that chunking relieved the time constraints they typically experienced when attempting to process real time speech in the target language (Chevillard, 1993).

There has been some speculation that dual processing -- that is, processing that involves both aural and visual modalities at once, such as television or film -- can derail comprehension of a single aural stream (Donaldson, 1976; Fisher, 1984; Singer, 1980). These multiple codes, it is argued, place too many demands on the processing capacity of the viewer, especially when the material is in a foreign or second language. MacWilliam (1986), for example, suggests that with multi modal processing there is a potential loss of information when it is presented via the aural channel accompanied by visual information of a non-linguistic nature. Students may consider the visual portion a distraction from the information delivered via the soundtrack. Where multiple

forms of input may hypothetically cause interference or cognitive/perceptual overload, studies involving second and foreign language students and subtitled video (a medium that combines visual, aural and textual elements) provide strong counterevidence. These combined media, on the contrary, appear to enrich both processing and recall of the target language. The addition of text does *not* interfere with comprehension (Borras, 1993; Chiquito, 1994; Jung, 1990; Vanderplank, 1990), but appears to stimulate deeper comprehension (Neuman et al, 1990), enhanced recall (Svensson and Borgarskila, 1985) and, consequently, more student output (Garza, 1991). Increased experience in multi modal viewing appears to improve the learner's ability to comprehend as strategies for optimizing processing input combinations are worked out over time (Salomon, 1979; Vanderplank, 1990).

Multimedia in Action

The following examines specific instances of multi modal processing as it assists the development of listening competency. Jack Richards (1985) describes listening competency as being comprised of a set of "microskills". These are the skills effective listeners employ when trying to make sense of aural input. A composite of Richard's thirty-three microskills required for conversational listening are presented along with corresponding scenarios¹ that demonstrate how multimedia technology can be called into service to support the development of each.

1. Retention of language chunks in short term memory.

Most current multimedia applications allow the student some control over the rate of language presentation. That is, users can start, stop, and review chunks of language to better understand and remember the aural text. The addition of video provides a clear, logical flow of events so that linking (remembering) new information to old is facilitated.

A learner of ESL is viewing a documentary on wildlife in a multimedia format. The voice of the narrator provides commentary on what is being visually depicted. This commentary is organized both in terms of the video on the screen and by standard narrative conventions in the aural text. The student moves from one short scene to another. When new information requires reference back to previously viewed material, the learner can click the mouse to review that old information and link it to the new.

2. Discriminate the sounds of the target language.

User control over language presented in more than one modality supports a student's ability to discriminate where words begin and end. The synchronized display of text along with the aural text assists the learner in distinguishing phonetic groupings and boundaries. When learners can see the faces of those speaking in the video, moreover, they can additionally make use of facial movements to understand the sound-meaning correspondence in the target language.

A learner of French as a foreign language is working with French television commercials in a multimedia format. On the screen she sees both the commercial itself and the verbatim text of the commercial's audio. The first time she listens to the initial ten-second chunk of the commercial, the actors' French sounds like a long, indistinguishable slur of words she does not recognize. By instantaneously repeating this sequence several times, by studying the accompanying text, and by relating what she is hearing and reading to the visual clues in the video, she works to hear, read, and see how these French sounds relate to actual words and their meanings.

3. Recognize patterns of stress, rhythm, and intonation and how they signal information and intent.

Stress, rhythm, and intonation are automatically highlighted when aural language is chunked into syntactic units (see Chunking). When we speak, the logical breaks in our discourse (the places where we pause ever so slightly) occur at syntactically predictable junctures. As such, when the aural text is chunked for the language learner -- when the presentation is paused at these junctures -- the learner begins to acquire a feel for patterns of sounds both rhythmic and syntactic. As far as how these patterns signal specific kinds of information and intent, the addition of video can carry this kind of information. When specific words are stressed and patterns of intonation used, learners can be cued to closely examine the visual and spoken reactions of interlocutors in the video presentation

A learner of German is working with a television mystery show in a multimedia format. Two persons -- the detective and a suspect -- are exchanging questions and answers on the screen. The detective is visibly growing more suspicious of the suspect. This suspicion is also reflected in the stress and intonation of his questions and remarks. Consequently, the suspect is becoming visibly more fearful. This is also reflected in the rise and fall of her speech. The learner is studying the effects of German prosodic contour. She can listen to each of the interlocutor's questions and answers as many times as and in any order she wishes while noting speakers' emotions and intentions.

4. Understand reduced speech.

Multimedia is particularly well suited to assist learners in their understanding of reduced forms of target language speech. Having the written version of fast, naturally-paced aural text on the computer screen allows the learner access to both the written and spoken forms simultaneously. That is, the learner may hear "wadjagonnado?" but will read "What are you going to do?" In this way, learners can come to understand the two different forms of the target language - spoken and written - as well as learn to decode these reduced forms.

An ESL learner is working with a parody of a U.S. comedy quiz show in a multimedia format. The quiz show contestants' quick, colloquial speech adds to the humor of the sequence. The learner hears a chunk of a contestant's speech -- "whaddjasay", and reads at the bottom of the screen "What did you say". He clicks on a control button on the lower part of the screen and a colorized version of the reduced form appears in contrast to the written form. The student can study the two text versions and repeat the aural version as many times as he likes. He also has the option of disabling the text entirely and testing himself on comprehending only the aural version of the reduced forms.

5. Recognize core vocabulary and the rules and patterns of words used to communicate.

Coordinated aural, visual, and textual information on the computer screen at the same time makes up an ideal laboratory for student problem-solving at the level of individual words and sentence structures. The learner has at her disposal rich visual and contextual clues that can assist in breaking the code of the written and aural text. The multi modal cues can be cross-referenced for word, sentence-level and broader understanding.

A student of Portuguese as a foreign language is working with a Portuguese movie in a multimedia format. She is just beginning to study the language and has a very limited vocabulary and knowledge of sentence structure. On the screen she sees young children arguing over a piece of fruit. From the visual circumstances, the verbatim text subtitles, in conjunction with the aural text, she is able to deduce which of the ten words on the screen refers to the pineapple. To double check, she clicks on a dictionary icon and types in the Portuguese word she thinks means pineapple. The definition is in Portuguese. She tries to read it, and makes out the cognate "fruta". She knows she's on the right track. She studies the other ten words in the subtitle, replays the sequence four more times, and observes how the word she now understands stands in relation to the other words in the sentence. She notices that one child is directly addressing the other and

surmises which of the words refers to "you" and that the verb is most likely in an imperative form. She goes on in a like manner to decode words and patterns in her new language.

6. Understand communicative functions of utterances according to context.

Video can be a very rich source of context for language processing. In a multimedia format, learners are provided control over the rate and order of video presentation and can therefore take advantage of starting and stopping the action in order to study language in a well represented context. Video also typically boasts tight correspondence between what is seen and what is heard. That is, in only very rare cases is the audio portion of video temporally disconnected to what is being viewed. By studying target language communication in a multimedia format, learners can experience and come to understand the connections between utterances and how they function within a visually depicted context.

A learner of Spanish as a foreign language is working with a television drama in a multimedia format. The action on the screen is between a man and a woman who are trying to decide what form of transportation to take. From the visual and aural context, the student perceives that the man's utterance "The bus takes too long" is communicating to the woman his unwillingness (his refusal) to agree to this mode of transport. When the woman replies, "Well, airline tickets are expensive" the learner checks the speaker's facial expression and intonation by repeating the sequence twice more, considers the nature of the characters' relationship as developed in earlier scenes, and concludes that the utterance is serving as an invitation to the man to offer to pay for their airline tickets. Recognition of communicative intent is enhanced by virtue of the multimedia format.

7. Process different speech styles, different rates, and performance errors.

Some multimedia software for language learning provides a slowed down version of the aural text. The learner can switch back and forth from a normal-paced text to a slower, sometimes simplified version of the target language audio. This is one approach. However, if a variety of video-based material is available in a multimedia format, a range of speech rate and styles is also likely, and less artificial. Rate and style of audio naturally vary according to the genre of the video selection. Many situation comedies, for example, exemplify slowed speech. Interlocutors speak slowly and deliberately so the joke can be processed and understood (Meskill, 1995). There are other kinds of programming, talk shows, for example, that are very fast-paced and difficult for non-native speakers to comprehend. Multimedia that includes varied genres

permits a broad experience of different voices with differing rates and speech styles. Students can control the aural text so they have sufficient time for their individual processing needs.

A first-year student of Russian is working with shows from Moscow Television in a multimedia format. The first sequence he views is of a news broadcast. The pace of the newscaster's speech is fast, but with the help of visual cues, the verbatim text, and an on-line Russian-English dictionary, he is able to make some sense of what the news story is about. He goes back to a menu of programs and selects a commercial. An old man in traditional Russian costume speaks slowly and carefully about his love for a particular brand of butter. Again, with the aid of the video and accompanying text, the student understands the message of the commercial. He has worked at comprehending two distinct rates of native speaker production.

8. Recognize that meanings can be expressed in different grammatical forms.

Redundancy in video presentations is common. That is, interlocutors and narrators frequently repeat the same information in different ways so that meaning and intention is made clear to the viewer. In a multimedia format, phrases and sentences that carry the same or similar meaning can be highlighted for users and/or the learner can be prompted to highlight those phrases and sentences she feels express like meanings. Highlighting can take the form of coloring text, visually juxtaposing two or three texts whose meaning is similar, or directing learners to click on portions of the text or video where they recognize redundancy.

A second-year student of Spanish as a foreign language is working with a Spanish documentary on European fashions in a multimedia format. The program combines voiceover narration and interviews with fashion models as they work. When the perspective switches from the narrator to an interviewee, the learner is prompted to note the narrator's text ("Most models worry about their weight") to that of the interviewee ("I am constantly dieting"). With both texts on the screen, the learner is guided to recognize the similarity in meaning of the two sentences. If she chooses, she may access a semantic grouping of additional words and phrases that express like ideas.

9. Infer meaning and make predictions using personal knowledge, experiences, and strategies.

Video is a medium to which language learners come well equipped. Students are very accustomed to inferring meaning and making predictions from what they see and hear on the screen. In a multimedia format, these viewing/comprehension strategies can be cued and guided

by, for example, posing pre-viewing questions on top of the stilled first frame of the sequence they are about to watch. Inference, predication, and calling up prior knowledge and experience can thus be activated.

A mid-beginning ESL learner is working with a short video sequence depicting a chef and her boss in a restaurant kitchen. Before clicking on the play button, three questions appear on the stilled initial frame of the video: "Who do you think these people are?" "Where are they?" "What do you think they will talk about?" The student studies the questions on the screen for some time and reflects on his answers. He then clicks on a notepad icon and types in "They are maybe cooks. They talk in a kitchen. They talk about food." He then plays the sequence, reviews his initial thoughts, and modifies them according to what he has understood from the actual scene.

In addition to these microskills for conversational listening, Richards also provides a taxonomy of microskills for academic listening (listening to lectures in formal instructional contexts). While it may seem desirable to provide students control and tools to work with recorded academic lectures, there are problems: first, having a talking head on the screen is a very poor use of video. It fails to take advantage of the medium's visual power. Second, the immediate, social/interactive nature of a lecture setting cannot be replicated. It is for these reasons that academic listening, as defined by Richards, is not treated in this discussion of multimedia and listening. The challenges of conversational listening can be more richly accommodated with this particular technology.

Listening Performance Activity

Brown (1994b) defines Listening Performance as what language learners actually *do* during a listening activity. In a multimedia environment, the possibilities for inviting students to do things before, during, and after listening in the target language are many and varied. Computer-based activity by nature implies an exchange between the user and the system. The user indicates something by typing on the keyboard, clicking a mouse, etc. The system responds to the user's input. This exchange is learner performance. In the realm of listening practice, this exchange can be limited to a learner indicating comprehension by selecting from multiple choice answers, or as interactive as a learner using information and tools supplied by the software to construct thoughtful representations of understandings of the aural text.

Brown defines six types of listening performance: reactive, intensive, responsive, selective, extensive, and interactive. These performance types progress from simple (reactive) to

the most complex form of listening (interactive). There is a parallel continuum when types of user performance with a multimedia system are examined. These range from simple reacting to on-screen or aural input (reactive) to complex interaction with material through manipulation and construction (interactive).

REACTIVE----INTENSIVE----RESPONSIVE----SELECTIVE----EXTENSIVE----INTERACTIVE

1. Reactive listening performance

This performance focuses on surface aspects of the language, not necessarily understanding. A prevalent example of this kind of listening is choral response listening where students directly mimic or perform simple transformations on what a teacher says. A multimedia example of reactive listening might be when a student is prompted to repeat or make transformations on what gets "spoken" by the system. This can be achieved by having students "speak" to the screen, type in what they hear, or click the mouse on a pictorial or textual representation of what is heard.

2. Intensive listening performance

Intensive performance requires learners to concentrate on the component parts of what they hear. This may take the form of listening to a teacher repeat a sentence and indicating the form of the verb or intonation pattern she is using. A multimedia example of intensive listening is colorizing. In a multimedia presentation where text accompanies audio and/or video, the user can be prompted to focus on component parts of sentences when these parts are visually marked by color. The user can also be prompted to listen for specific components and type them in or colorize them to indicate successful discrimination.

3. Responsive listening performance

This performance requires students to listen to a teacher's question or cue and respond immediately and appropriately, thereby indicating understanding. Teacher prompts can take the form of *meaningful questions* (Where were you yesterday?), *commands* (Please close that door.),

clarifications (What did she say?), and *comprehension checks* (Do you mean she was sick?). In terms of *meaningful questions*, multimedia systems are limited. Computers cannot process nor respond to natural input. However, the system can certainly pose such questions, prompt the student to type in or record a response, and save these as files for a peer or teacher to assess. On the other hand, responding to *commands* is a performance perfectly suited for multimedia systems. The learner can respond to audio commands in any number of ways: clicking on the screen, moving objects around on the screen, starting and stopping an audio or video segment as commanded, typing in predetermined words or sentences, and the like. Responding to requests for *clarification* is also feasible in a multimedia environment. Requests for clarification (Are you sure? Do you mean X?, etc.) can be simulated auditorily, textually, and/or by a video character. The learner has only to provide clarification in ways similar to those suggested for responding to commands. Demonstrating *comprehension* is also readily feasible with this form of instructional technology. Comprehension can be indicated in response to multimedia prompts through typing or clicking with corresponding feedback provided.

4. Selective listening performance

This performance requires learners to listen to longer stretches of discourse for the purpose of getting specific information from the aural text. Multimedia also accommodates this listening performance well and easily. Learners can be prompted to listen selectively for particular information, then to indicate successful identification of this information by typing in or selecting appropriate key words, pictures, or sequences from a group of possible selections. The learner can also manipulate elements on the screen in response to successful selective listening.

5. Extensive listening performance

Extensive listening requires fuller understanding of lengthier aural texts for the purpose of in-depth understanding. This form of listening is especially well suited for multimedia in that the learner, unlike in real-time situations, can control the rate and sequence of the aural presentation. She can also make use of visual and textual clues available in a multimedia format to understand what she hears. On-line notetaking capabilities, access to supporting information, and the availability of tools (key word guidance, a dictionary, and the like) add to the suitability of the medium for extensive listening. Learner performance can take the forms described in previous sections, only require more in-depth understanding.

10. Interactive listening performance

Interactive performance calls into play the above types of listening performance in face-to-face interaction. As the full negotiation of meaning that takes place between human interlocutors is not realizable between a learner and a multimedia system, the option of using the technology as a springboard for student-student interaction becomes an alternative (see Pairwork with Multimedia below). Individual work with listening skills development can, moreover, be viewed as needed rehearsal for human interaction in the target language.

-Caveats

1. Congruence

If including visuals is to advantage the language learner's listening skills development, certain conditions need apply. For example, tight correspondence between visual and aural elements in video is more likely to increase comprehensibility than would incongruence. Direct, tangible correspondence between what is uttered and what the utterance refers to is a fundamental characteristic of aural input that is readily comprehensible (Dulay et al, 1982). Environmental, kinesthetic, and non-verbal messages in the form of human gesture and movement can supply supporting cues for learners in decoding aural messages. Paralinguistic cues not only contribute to understanding the surface meaning of utterances, but also provide more subtle information such as cues to the speakers' intent (Garza, 1991; Kelly, 1985, Riley, 1979). It is important, therefore, that the visual and the aural cohere and be mutually supporting in representing meaning. It is a natural human reaction to work very hard at making sense of things that are oddly or unpredictably juxtaposed. However, when there is incongruence between the two channels, a language learner must expend energies on forcing an interpretation, rather than simply decoding the message. In the case of learning language where the focus is on strategizing the comprehension of an aural text, the visual can be an asset to learner comprehension when it is *aligned with* that text.

2. Familiarity

In addition to correspondence between visual and aural representation, the association of visual images with the meaning they represent is more likely when the visuals are salient to students' cultural experiences (Walker de Felix et al, 1990). In a study with U.S. students learning Spanish, for example, the television show *Sesame Street* was shown in the target language. Learners reported using scenes, characters and themes that were already familiar to them to make sense of the program. Subjects recounted that familiarity with elements of the program directly assisted their comprehension (Pearson, 1978). In addition to calibrating the match

between the aural and the visual, consideration should also be given to cultural salience

3. Integration

Integration of any technology into the larger context of learning requires correspondence of goals and content between the two realms (Meskill & Shea, 1994). If what gets learned and practiced using multimedia is closely aligned with and recognized by other learning activities that take place in other learning contexts, success is more likely. If there is recycling and follow-through on content, success is also more likely. One example of valuing listening skills practice with media in other contexts is the use of supporting, off-line materials. Availability of such materials have been found to be critical to technology-mediated language instruction. For example, in a 1983 study of televised language programming, Lo found that significant improvement in foreign language skills development can only be achieved when there is extensive support materials (e.g., print and audio) that are closely keyed to what happens on the television screen (Lo, 1983).

4. Pair Interaction with Multimedia

In the era of communicative language teaching and learning, primary concern is given the development of the learner's ability to actively negotiate meaning in the target language. It is through processes involved in two-way communication that the rules and structures of the target language become incorporated into the learner's L2 system. Moreover, the depth of instructional experiences increases when involvement with another is part of the process (Pica & Doughty, 1985; Stevik, 1976). A disadvantage of learning technologies is that interaction is limited to machine prompts and reactions. The machine-based conversation consequently lacks the multiple and complex elements of human interaction that contribute to negotiated meaning and, ultimately, the development of communicative competence. Adding this missing element by pairing learners at the computer is an approach that may add the dimension of actively negotiated conversation.

By pairing learners, the fact that images are interpreted differently by individuals, for example, can be capitalized on in the language classroom and interpretive skills and processes exploited (Jiang & Meskill, 1995; MacWilliam, 1986). As regards the interpretative value inherent in a medium such as video, pairing students to co-view makes sense (Walker de Felix et al, 1990). Differing perspectives on what happens on the computer screen can provoke interchange between students that may carry some pedagogical value; e.g., practice in face-to-face communication in the target language. Active co-viewing and conversation with multimedia playing the role of catalyst seems an attractive pedagogical approach. Not only can

students develop listening skills by directly controlling the technology, but they can also benefit from *negotiated* discourse processes with their partners.

Although the notion of computer prompted discourse between students holds intuitive appeal, the majority of studies that have examined interaction between paired language learners to date portray something quite different. Where pair and groupwork with native speakers seems to hold some promise, when the common language is a second or foreign one, conversational miscues and breakdowns are frequent (Abraham & Liou, 1991; Chang & Smith, 1991; Legenhausen & Wolff, 1990; Levy & Hinckfuss, 1990; Meskill, 1993; Mydlarski, 1987). One reason for this may be that conversation at the computer, especially in conjunction with the demands of multi modal processing, places linguistic demands on participants that can derail sustained, involved interaction. Language learners do not necessarily possess the linguistic tools needed to keep a conversation going while attending to a technology that demands multi modal processing. Moreover, activities like watching video and computing are typically solitary activities that preclude sustained, involved interaction between individuals. The mindset for each of these media-based activities does not accommodate a human interlocutor. Pairwork with multimedia therefore may not be particularly supportive of listening skills development due to the lack of sustained, involved interaction between paired learners, and due to the potential processing overload when a "third party" is introduced into the configuration (Meskill, 1992). Pairing students with multimedia materials in the hopes they will engage in rich exchanges, then, needs to be approached with some caution. Students need models for and guidance in this type of three-way conversation with combined media.

Arguments in support of multi modal processing suggest that multimedia can serve as a powerful tool for an individual's listening skills development. The forms visual and aural material take and how these are keyed and supported within the instructional environment are, of course, critical. Individual interaction with aural, visual, and textual information can serve the learner's needs up to a point, but cannot provide opportunity for fully negotiated interaction. Pairing learners with the goal of their conversing, while carrying some intuitive appeal, does not insure that negotiated discourse and accompanying listening practice will result.

CONCLUSION

The case for multimedia as a technology that supports listening skills development in another language is strong. Research and anecdotal accounts, observations of users, and attitudinal feedback indicate that the medium motivates target language processing in general

and multi modal processing in particular. Arguments supportive of multi modal processing as a means of listening skills development emphasize: 1) the role of text and visuals as aids to language processing when appearing in conjunction with the aural text; 2) the motivational aspect of video as an advantage for language instruction; 3) the fact that combined media enrich target language processing, thereby rendering input more direct and salient for the language acquisition process.

Clearly, no technology can replicate the linguistic growth derived from human interaction (Ur, 1984). Multimedia technology can only simulate a very limited conversation: meaning gets only partially negotiated - partial negotiation being a one-way effort on the part of the individual user. The learner is consequently limited to the role of an "overhearer" (Rost, 1993) or eavesdropper, rather than that of a participant. The full, two-way active negotiation of meaning considered essential to successful language learning is simply not possible via computer and the prospects for needed natural language parsing remain dim (Dreyfus and Dreyfus, 1990). Multimedia-based practice in listening skills development can, however, be viewed as rehearsal for face-to-face interaction with multi modal processing contributing to the larger L2 acquisition process.

As a medium for learning language, multimedia represents a myriad of instructional possibilities. As a tool for listening skills development, there is a logical match of system characteristics (combining text, audio and video) and the goal of listening skills development in a second or foreign language. Careful consideration on the part of teachers and software developers of the range of possibilities for combining input modalities and tools that empower student manipulation of them is essential.

REFERENCES

- Abraham, R. and Liou, H. (1991) Interaction Generated by Three Computer Programs. in Dunkle, P. (Ed.) Computer-Assisted Language Learning and Testing: Research Issues and Practices. New York: Newbury House.
- Baggett, P. (1979) Structurally equivalent stories in movie and text and the effect of the medium on recall. Journal of Verbal Learning and Verbal Behavior, 18, 357-354.
- Borras, I. (1993) Effects of subtitling on speaking performance of college students of French learning with multimedia courseware. Paper presented at the 1993 Computer Assisted Language Learning Consortium (CALICO), Williamsburg, VA.
- Breen, M. & Candlin, C. (1980) The essentials of a communicative curriculum in language teaching. Applied Linguistics, 1.
- Brinton, D. & Gaskill, W. (1978) Using news broadcasts in the ESL/EFL classroom. TESOL Quarterly, 12, 4, 403-413.
- Brooks, Z., Bush, M., Meskill, C., Marlino, M. Reeves, T. and Verano, M. (1990) Project Communication: Interactive videodisc second language instruction in high schools. Final Report, Department of Education.
- Brown, H.D. (1994a) Principles of Language Learning and Teaching. New York: Prentice-Hall Regents.
- Brown, H.D. (1994b) Teaching by Principles: An interactive Approach to Language Pedagogy. New York. New York: Prentice-Hall Regents.
- Brownfield, S. (1990) Project Athena Alpha Evaluation: French Videodisc. Report of the Center for Educational Computer Initiatives, Massachusetts Institute of Technology.
- Call, M. (1985) Auditory short-term memory, listening comprehension, and the input hypothesis. TESOL Quarterly, 19: 765-81.
- Chang, R. and Smith, F. (1991) Cooperative learning and CALL/IVD in beginning Spanish: an experiment. The Modern Language Journal, 75 (2) 205-211.
- Chevillard, I. (1993) Study of a multimedia application for French. Report. University at Albany, Center for Electronic Language Learning and Research.
- Chiquito, A. (1994) Metacognitive learning techniques as part of the user-interface: An approach to interactive applications' design or foreign language learning. Unpublished manuscript, Center for Educational Computer Initiatives, Massachusetts Institute of Technology.
- Collins, W. (1981) Schemata for understanding television. in Kelly & Gardner (Eds) Viewing Children through Television. San Francisco, CA: Jossey-Bass.
- Diller, K. (1981) Neurolinguistic clues to the essentials of a good language teaching methodology: Comprehension, problem solving, and meaningful practice. in H. Winitz (Ed) The comprehension approach to foreign language instruction. 141-153. Rowley, MA: Newbury House.

- Donaldson, M. (1976) Children's Minds. London: Fontana.
- Dulay, H., Burt, M. and Krashen, S. (1982) Language Two. New York: Oxford U. Press.
- Dunkle, P. (1991) Listening in the native and second/foreign language: Toward an integration of research and practice. TESOL Quarterly, 25, 3,431-457.
- Dunkle, P. (1986) Developing listening fluency in L2: Theoretical principles and pedagogical considerations. Modern Language Journal, 70, 2, 99-106.
- Dreyfus, H. and Dreyfus, S. (1986) Mind over Machine: The power of Human Intuition and Expertise in the Era of the Computer. New York: Free Press.
- Fisher, Eunice (1984) Television and language development. Journal of Educational Television, 10, 2, 85-90.
- Garza, T. (1991) Evaluating the use of captioned video materials in advanced foreign language learning. Foreign Language Annals, 24, 3, 239-258.
- Gay, G. (1986) Interaction of learner control and prior understanding in computer-assisted video instruction. Journal of Educational Psychology. 78, 3, 225-227.
- Hawkins, R., Yong-ho, K. and Pingree, S. (1991) The ups and downs of attention to television. Communication Research, 18, 1, 53-76.
- Jiang, M. and Meskill, C. (1994) Multimedia ToolBook for Teaching English as a Second Language. in Sechrest, T., Thomas, M. & Estes, N. (Eds) Leadership for Creating Educational Change: Integrating the Power of Technology. Austin, TX: Morgan Printing.
- Jonassen, D. (1993) Conceptual frontiers in hypermedia environments for learning. Journal of Multimedia and Hypermedia, 2, 4, 331-335.
- Jung, U. (1990) The challenge of broadcast videotex to applied linguistics. IRAL, 28, 3, 201-220.
- Kelly, R. (1985) Video and language learning: An approach to viewing comprehension. RELC Journal, 16, 1, 52-57.
- Koskinen, P. and Gambrell, L. (1993) Using captioned television to enhance the vocabulary of adult beginning readers. Paper presented at the American Educational Research Association, Atlanta, GA.
- Krashen, S. (1994) The pleasure hypothesis. Georgetown University Round Table on Languages and Linguistics 1994. Washington DC: Georgetown U. Press.
- Krashen, S. (1980) The input hypothesis. In J. Alatis (ed) Current Issues in Bilingual Education. Washington, DC: Georgetown U. Press.
- Lambert, W. (1986) "Pairing First- and Second- Language Speech and Writing in Ways that Aid Language Acquisition", in Language Processing in Bilinguals: Psychological and Neurological Perspectives, J. Vaid (ed) Hillsdale NJ: Erlbaum.

- Legenhausen, L. and Wolff, D. (1990) CALL in use-use of CALL: Evaluating CALL Software. System, 18, 1-13.
- Levy, M. and Hinckfuss, J. (1990) Program design and student talk at computers. Computer Assisted English Language Learning Journal, 1, 21-26.
- Lo, J. (1983) Research: Video in language teaching. Media in Education and Development, 16, 3, 147-149.
- Luke, C. (1985) Television discourse processing: A schema theoretical approach. Communication Education, 34, 91-105.
- MacWilliam, I. (1986) Video and language comprehension. English Language Teaching Journal, 40, 2, 131-135.
- Markham, P. (1989) The effects of captioned television videotapes on the listening comprehension of beginning, intermediate, and advanced ESL students Educational Technology, 29, 38-41.
- McCarthy, M. (1991) Discourse Analysis for Language Teachers. New York: Cambridge U Press.
- Meskill, C. (1995) Commercial television and the learner of English as a second language: Implications for linguistic and cultural development. in K. Swan, C. Meskill & S. Dimeo (Eds) Social Learning from Television. Cresskill, NJ: Hampton Press.
- Meskill, C. & Shea, P. (1994) Multimedia and language learning: Integrating the technology into existing curricula. Proceedings of the Third Conference on Instructional Technologies. State University of New York, FACT.
- Meskill, C. (1993) ESL and multimedia: A study of the dynamics of paired student discourse. System, 21, 3, 323-341.
- Meskill, C. (1992) Off-screen talk and CALL: Role of the machine/participant Computer-Assisted English Language Learning Journal. 3, 1, 1-9.
- Meskill, C. (1991a) IVD for English as a second language: The anatomy of repurposing. Journal of Interactive Instruction Development, 3, 4, 26-29.
- Meskill, C. (1991b) Multimedia and language learning: Assessing goals and system attributes. Computer-Assisted English Language Learning Journal. 2, 2, 11-14.
- Meskill, C. (1991c) Language learning strategies advice: A study on the effects of on-line messaging. System. 19, 3, 277-287.
- Mueller, G. (1980) Visual contextual cues and listening comprehension: an experiment. Modern Language Journal, 64, 3, 335-340.
- Mydlarski, D. (1987) Cooperative computer-assisted language learning: Is it living up to its promise? Journal of Educational Techniques and Technologies, 20, 26-29.

- Neuman, S., Burden, D. and Holden, E. (1990) Enhancing children's comprehension of a televised story through previewing. Journal of Educational Research, 83, 258-265.
- O'Malley, M., Chamot, A., Stwener-Manzanares, G., Russo, R. and Kupper, L. (1985) Learning Strategy Applications with Students of English as a Second Language. TESOL Quarterly, 19,3, 557-584.
- O'Malley, J., Chamot, A. and Kupper, L. (1989) Listening comprehension strategies in second language acquisition. Applied Linguistics, 10, 4, 418-437.
- Omaggio, A. (1979) Pictures and second language comprehension: do they help? Foreign Language Annals 12, 107-16.
- Oxford, R., Crookall, D, Cohen, A., Lavine, R., Nyikos, M., Sutter, W. (1990) Strategy Training for Language Learners: Six Situational Case Studies and a Training Model. Foreign Language Annals, 22, 3, 197-216.
- Pavio, A. (1965) Abstractions, imagery, and meaningfulness in paired associated learning. Journal of Verbal Learning and Verbal Behavior, 10, 1-12.
- Pearson, S. (1978) *Plaza Sesamo* and Spanish Language Learning. ERIC Document Number ED155890.
- Pederson, K. (1988) Research on CALL. in W. Smith (ed) Modern Media in Foreign Language Education: Theory and implementation. Lincolnwood, IL: National Textbook.
- Pica, T. and Doughty, C. (1985) Input and interaction in the communicative language classroom: a comparison of teacher-fronted and group activities. in S. Gass and C. Madden (eds) Input in Second Language Acquisition, Rowley, MA: Newbury House.
- Price, K. (1983) Closed-captioned TV: An untapped resource. MATSOL Newsletter 12, 2: 1, 8.
- Richards, J. (1985) The Context of Language Teaching. New York: Cambridge U. Press.
- Riley, P. (1979) Viewing Comprehension "L'oeil ecoute." Melanges pedagogiques. 10.
- Rost, M. (1993) Listening in Language Learning. New York: Longman.
- Salomon, G. (1979) Interaction of Media, Cognition and Learning. Washington DC: Jossey-Bass.
- Salomon, G. and Leigh, T. (1984) Predispositions about learning from print and television. Journal of Communication, Spring.
- Savignon, S. (1991) Communicative language teaching: State of the art. TESOL Quarterly, 25, 261-277.
- Secules, T., Herron, C. and Tomasello, M. (1992) The effect of video context on foreign language learning. The Modern Language Journal, 76, iv: 480-490.

- Singer, J. (1980) The power and limitations of television: A cognitive-affective analysis. The Entertainment Functions of Television. P. Tannenbaum (Ed) Hillsdale, NJ: Erlbaum.
- Snyder, H. and Colon, I. (1988) Foreign language acquisition and audio-visual aids. Foreign Language Annals, 21, 4, 343-348.
- Stevens, V. (1989) A direction for CALL: from Behaviorist to Humanistic courseware. in M. Pennington (ed) Teaching Languages with Computers: The State of the Art. La Jolla, CA: Athelstan.
- Stevick, E. (1976) Memory, Meaning and Method. Rowley, MA: Newbury House.
- Svensson, S. & Borgarskola, M. (1985) Video, authenticity, and language for special purposes teaching. Foreign Language Annals, 18, 2, 149-152.
- Tudor, I. (1987) Video as a means of cultural familiarization. System, 15,2, 203-207
- Underwood, J. (1990) Research in Hypertext: Desiderata. Computer Assisted English Language Learning Journal, 1, 4: 33-36.
- Ur, P. (1984) Teaching Listening Comprehension. Cambridge: Cambridge U. Press.
- Vanderplank, R. (1990) Paying attention to the words: practical and theoretical problems in watching television programmes with uni-lingual (CEEFAX) sub-titles. System, 18, 2, 221-234.
- Walker de Felix, J., Johnson, R. and Schick, J. (1990) Socio- and psycholinguistic considerations in interactive video instruction for limited English proficient students. Computers in the Schools, 7, 1, 173-190.

¹The scenarios of language learners using multimedia are based on actual observations and user feedback culled from sessions in the Center for Electronic Language Learning and Research, University at Albany, 1992-1995. Students working in the Center use "repurposed" videodisc sequences. These are commercially available videodiscs that have been refitted with computer templates for language learning purposes. (For a full description of repurposing, see Meskill, 1991a)