RECENT TRENDS IN LANGUAGE TESTING: THE CASE OF TESTING ORAL LANGUAGE

Heidi Byrnes
Georgetown University

Before one can engage in language testing in general, in testing a learner's speaking ability specifically, one must first clarify a number of central issues regarding essential features of language and its use and how these are acquired by non-native speakers.

For this reason I will address the following major points. I will begin by exploring the connection between testing and our basic professional beliefs. I will then summarize current assumptions in foreign language pedagogy that have attained prominence in the United States and Europe, assumptions that are captured by the term "communicative language teaching." The unique challenges to established testing practice brought on by this shift will be outlined in the third section. And finally, the largest portion of this paper will be devoted to exploring approaches to oral language testing that have developed in the United States in the last decade or so.

It is worth emphasizing that these steps are far from complete and by no means universally accepted. In addition, a cautionary note is appropriate: their presuppositions about language and language teaching should not be transferred uncritically, but should be examined carefully in order to determine whether and how they might apply to the Japanese situation. Only then can the American experience provide input that leads to viable solutions.

LANGUAGE TESTS AND OUR PROFESSIONAL BELIEFS

Any testing is inherently a sampling procedure. Tests cannot possibly include everything that was taught prior to the test, nor can they project precisely how learners will have to use their language skills after the test. As a consequence the test developer must choose, ideally selecting those aspects of language which are important. But "importance" does not exist of its own. Rather, it comes about as a result of numerous considerations.

The most important of these considerations regarding "importance" pertains to

Heidi Byrnes is a Professor of Linguistics at Georgetown University.

the goals of language teaching and learning which apply in a given case. It follows that if speaking ability is to be tested, then speaking ability must have been targeted as a goal of instruction, and instruction must, in fact, have allowed students to engage in spoken interaction in order to acquire this complex ability in a well-articulated and well-motivated curricular sequence. While this may sound all too basic to be worthy of repetition, there is every indication that many programs, be they at the institutional, state, or even national level, have major gaps precisely in this area. Numerous mismatches exist. For example, one teaches something in a way that really does not support the stated long-term goals; one tests abilities that have not been explicitly taught; or one tests something that has been taught, but which bears little resemblance to what learners ultimately are expected to do with the language.

Thus, only if speaking ability is the explicit goal of a program and if the instructional approach has provided opportunities for the learners to develop it, is it appropriate and fair to test this multi-faceted skill.

In addition to the mismatches between teaching and testing just mentioned, both teaching and testing have tended to focus on short-term goals, usually the more easily defined, more form-related components of language which, inherently, are more amenable to prevailing modes of testing. If it is true that instruction rarely "gets to" or "has time for" enhancing the learners' acquisition of the long-term, comprehensive functional goals, then testing practice is even more remiss in addressing these aspects which are much more descriptive of what using a language is all about.

While reference to overall goals, including linguistic, pragmatic, discourse, and sociolinguistic competence, is critical for language testing (see Canale and Swain, 1980), "importance" comes about also by considering aspects of second language

acquisition. What steps are likely to be necessary for a learner to attain the global functional ability to which she or he aspires, or, rephrased in terms of testing, what features of the language should the learner have acquired at what stage with what degree of certainty to allow us to feel confident that she or he is progressing satisfactorily?

Second language acquisition research obviously admits to many lacunae regarding how language is learned. Even so, numerous studies have convinced us, particularly through the concept of interlanguage, that such learning is not a simple on/off procedure, such that students can produce a form correctly all the time or they cannot. Instead, learning is a lengthy and, at times, circuitous road, leading from initial awareness of forms and their meaning, to their better understanding, to halting and error-prone variant use of these forms in restricted contexts, and, finally, to complete mastery in all contexts (for studies in interlanguage development, see Eisenstein, 1989).

If one applies these insights to the development of speaking ability, one must conclude that neither past teaching nor past testing have adequately reflected the fact that correct speaking in all contexts is the very last step in a long process. Thus, testing will have to recognize intermediate stages which indicate that learners are progressing steadily toward this lofty and demanding goal.

While the concern with intermediate stages and goals of learning extends to all aspects of use, the form side of language is likely to be particularly prominent in instructed language learning. The question then becomes: how do we test learners' command of specific language forms while recognizing that these forms are a means to an end, namely successful communication, rather than the end itself?

It is impossible in this context to provide details, if for no other reason than that they

Heidi Byrnes

are likely to be language-specific. But one overall observation is appropriate and critical: the formal components of a language carry different weight at different points in the process of learning, and this differentiation must be incorporated into testing practice.

For example, without a doubt learners must learn to produce proper past tense forms for English irregular verbs, such as "to go-went-gone", and, therefore, one may wish to test them on this knowledge. However, in isolation these forms and the degree of control a learner has over them hold no inherent significance. Only in terms of their place and role in language learning and language use do they become important, an observation which implies that the "importance" or role of a given form shifts over time. To return to the earlier example, once students possess basic familiarity with English irregular verbs, it makes little sense to test them as decontextualized forms in terms of mastery. Their real use and usefulness, and thus their real importance, lies in enabling learners to provide extended narratives. Such narration has certain textual requirements, among them devices of coherence and cohesion typical for the English language, alongside proper verb morphology. Thus, in terms of the learners' ability to narrate, it is less critical that they can provide the form "went" correctly in a list of irregular verbs and more informative of their progress if they can create utterances such as "Afterwards we went to a restaurant where we enjoyed a pleasant meal and wonderful conversation. However, unfortunately, my sister could not come along but went home because of a terrible headache."

I have argued that we can assign importance to specific features of learner performance, and thus devise appropriate testing modes and procedures only on the basis of stated goals and by considering pervasive aspects of language learning.

Given this primacy for goals articulation it is all the more surprising that this area frequently receives only perfunctory attention in foreign language pedagogy. To avoid possible misunderstanding, by no means is it the case that all programs should aspire to the same goals in lock-step fashion, whether that is communicative ability or language learning for the sake of enhancing one's research access. By the same token, it is also inadmissible to simply drift along without having clarified objectives and ways of reaching them. The following parameters would seem to apply in this critical decision-making process:

• Language teaching and learning, and by implication language testing, is embedded in social contexts. What social contexts might impact on the setting of goals? For example, knowing another language may be considered to be a characteristic of the educated elites. If these educated elites have little occasion or little incentive for personal contact with speakers of other languages, and if the overall cultural climate is toward internal self-sufficiency rather than reaching out to others, then the capability to read foreign language texts, literary or non-literary, is a valid goal. A commonly used indicator to gauge comprehension is translation of texts into the native language. Intermediate goals toward attaining this ability would include an extensive vocabulary that deals with issues in the target language culture, augmented by facility with dictionaries, extensive familiarity with the literate norms of the language, and the ability to analyze texts for their literary value. By contrast, good pronunciation habits, fluency in production, ready access to the vocabulary of daily life, or familiarity with the interactive norms pertaining to speaking, would hardly be of interest, neither in teaching nor in testing.

• Aside from social expectations, goals reflect a network of professional expectations and convictions. Not infrequently these two clash. For instance, in the United States

the senior university professorate in language departments continues to consider the study of literature to be of paramount importance and makes demands of students similar to the case just described. By comparison, the dominant pedagogical discussion takes language as a means for interactive communication, as a way of performing socially derived tasks, of give and take in speaking and, increasingly more, also in reading and writing. Obviously, such divergent attitudes and goals for language learning surface not only in teaching, they also affect testing practice.

• Finally, and certainly not least importantly, the individual learner comes to the task of language learning with certain explicit or implicit goals. As recent research into motivation, attitudes, and the role of anxiety in language learning has brought to the forefront again (Horwitz and Young, 1991), it is the individual who learns or does not learn a language because she or he feels that instruction does or does not address what she or he wants to accomplish. An emphasis on the individual learner does not negate the fact that instructional systems inherently group together people with a range of goals, where some of these may not even agree with a particular institution's mission. But it is best to uncover that information at the outset, to share it with learners, and to adjust the instructional approach accordingly. Among the alternatives are the attempt to have learners modify their expectations in the direction of those underlying a given program, or, in reverse, of altering institutional goals so that they can reflect what may well have become general social trends regarding the purposes of language learning. For example, if trends in student interest in the United States bear any resemblance at all to developments in Japan, then the demands for communicative language teaching that reaches out into the professions must be taken seriously.

To conclude this section on the connection between testing and our professional beliefs, let us focus on the relationship between testing and curricular goals. Under ideal circumstances one first determines overall goals, spells these out as a sequence of intermediate and specific goals, and then expresses them in terms of learner performance statements. Only then can and should testing procedures be devised. But we may have, in foreign language pedagogy, a situation that is the exact reverse of what things ought to be. Tests often are the way they are because it is easiest to test in a certain fashion, not because these testing procedures and the resulting test items reflect our goals.

However, problems do not stop here. Any language teacher has the desire to appear as having achieved what she or he set out to achieve. We want to be successful and we want our students to be successful. One of the easiest ways for creating at least the appearance of success is to direct our teaching toward our testing. In other words, tests that were originally devised because of considerations of testing expediency or for psychometric reasons, all of a sudden drive our teaching practice and ultimately our curriculum.

One all too obvious example shall suffice. We know that tests that require students to perform all kinds of manipulative tasks, such as switching nouns from singular to plural, altering the subject, or changing the tense of a sentence, can be created relatively easily and can also be checked and scored without too much difficulty. We also know that, necessary though these skills may be, their relationship to functional language use in speaking is quite tenuous. How many students with commendable scores on tests made up of such items utterly lack an ability to communicate anything in speaking? Yet, despite this repeated experience, disturbingly little has changed in our approach to testing language performance, even when speaking ability is explicitly stated to be the instructional goal. On the contrary, the practice of discrete-point, decontextualized testing has created

the practice of discrete-point decontextualized teaching. The amount of time students devote to memorizing lists of vocabulary or forms, just so that they will do well on the test, leaving little time for using language in context, is a stark reminder of the sway testing practice can hold over teaching.

Of course, it is difficult to say whether this practice shows that old teaching goals, with their emphasis on formal accuracy in isolation, have not changed much, in which case such testing would, in fact, be appropriate, or whether old testing traditions continue to affect classroom practice, even when that classroom now targets functional ability in the language, in which case such testing is diametrically opposed to the goals of teaching. Or perhaps, we have a mixture of the two, both in the American as well as the Japanese context. Whatever the precise configuration, our societies increasingly demand functional use of the language as the outcome of language learning; as a profession we have espoused communicative language teaching; and the majority of learners thinks of language much more as a tool to be used in various professions, than as a subject that primarily leads to familiarity with Literature. The task for us is to find suitable testing procedures that reflect and enhance our teaching efforts.

COMMUNICATIVE LANGUAGE TEACHING AND LANGUAGE LEARNING

If communicative language use is, indeed, a valid instructional goal, we must first be aware of its critical features so that, ultimately, we can incorporate them in our testing procedures. What, then, characterizes such communicative language teaching and learning?

The Relationship of Norm and Variation

Perhaps the greatest change results from our expanded understanding of the systemicness of language, and by implication our tolerance of, indeed, inclusion of variation in language and language learning.

Our previous one-dimensional interpretation of "system" dealt with language only as a formal entity. However, once performance became a focus for inquiry, this unidimensional understanding had to give way to a multidimensional understanding which extends the concept of system to situationally and socially determined aspects of language use. As a consequence, we cannot be content with teaching a single language norm that consists of correct forms. We must teach a variable system, which recognizes the social and situational context within which language is normally used. True, classrooms have inherent limitations in that regard. However, they also have possibilities, many of which have hardly been explored.

The Role of the Individual Learner

Inextricably, a shift from language as a system of forms to a system in use involves a dramatic shift in the role of the individual learner. From the innumerable ways for describing this shift, perhaps none is more critical than the learner's social role, the level of involvement and responsibility in the act of learning which students must now take on. As long as mastery of language forms is the goal, learners cannot actively contribute much to the enterprise. Vocabulary and syntax are generally either right or wrong, with not a lot of choice or variation. However, as soon as use is the long-term instructional goal, choices must be made. There is not only one correct way to apologize in a language. There are many, and these are socially and situationally specified and chosen individually by the speaker.

All of this amounts to a central requirement for the new communicatively oriented foreign language class: the individual must be given greater freedoms and,

commensurately, must accept greater responsibilities in the process of learning. This leads directly to the following point.

Language Replication and Language Creation: Second Language Norm and Interlanguage

One can say that, up until now, the learners' central task was to replicate the model of the instructor. Imitation to the best of their ability would summarize their level of engagement. Now, however, they are asked to create language and to make choices, something that has complicated their task tremendously (Swaffar, 1989). Small wonder that, instead of producing correct forms in a mastery mode, they are likely to offer more or less successful approximations of the target language system, a phenomenon that is captured with the term "interlanguage."

This term is not simply a new way of referring to the old troublesome occurrence of errors. Far different, the concept of interlanguage recognizes that language learning, as a process of approximations toward the multiple norms of a language, is inherently error-prone. Testing with a communicative orientation must recognize this fact and attempt to incorporate it in defensible ways into testing practice.

Comprehensible Input - Comprehensible Output

In contrast with wide-spread misperceptions, communicative language teaching is by no means the equivalent of a near-exclusionary emphasis on speaking. On the contrary, as previously mentioned, it recognizes the critical relationship between comprehension and production. "Comprehensible input", a term popularized by Krashen, thus became one of the key terms in the initial stages of a move toward communicative language teaching (Krashen and Terrell, 1983). However, gradually, the notion of output as being critical for the development

of language skills has provided a much-needed balance (Swain, 1985). At the very least, a pedagogy which recognizes the special characteristics of input processing as well as output processing will differentiate between language use that primarily relies on the learner's background knowledge and is semantically based (reading and listening), and language use that must also focus on specific language forms, is thus syntactically based (speaking and writing). It goes without saying that, for most second language learners, comprehension tasks can be at a significantly higher level than production tasks.

If one compares such an approach with much of current classroom practice one observes a curious mismatch. In the beginning, students often get simplistic comprehension tasks in reading and in listening, while they are expected to perform tasks in speaking that go far beyond their capabilities. By contrast, in more advanced classes, students are frequently asked to read anything printed in the foreign language. But their speaking involves little more than single sentence answers, or even just the completion of the teacher's sentence with a phrase or a word. After all, the instructor knows all too well that an appropriate discussion of a given text would go far beyond the learners' abilities. The result is a high level of frustration due to a total reversal of the natural relationship between receptive and productive skills and their development. Instead of initially incorporating students' background knowledge to compensate for their limited knowledge of the language, we present them with rather unenticing texts that often come close to insulting their intelligence. At the same time beginning language classes often demand a tremendous amount of processing in speech right from the start. Later on, when learners have begun to automatize certain aspects of language in their speech, instruction often does not allow them to grow. Essentially we lack a pedagogy which supports the development of discourse competence and,

therefore, students continue to be limited to short phrases or even individual words. In terms of both issues, teaching and testing for oral proficiency, a major reorientation seems in order.

The Role of the Teacher

Though any new role for the learner inevitably affects the teacher, one particular aspect of the teacher's role in communicative language teaching deserves special attention. In the new language creation paradigm teachers are not so much impeccable models, policemen over accuracy at all cost, as they are facilitators of student-directed, creative learning. A much higher level of knowledge regarding the process of language acquisition is now required of them: aside from linguistic knowledge, teachers must impart to their students discourse, pragmatic, and sociolinguistic knowledge. Also, the teacher has become the best-informed analyst of each individual student's progress, one who can provide well-sequenced opportunities for additional growth. One might say that the teacher continues to direct all learning, only this time, as it were, less visibly from behind the scene.

THE CHALLENGES OF LANGUAGE TESTING UNDER A LANGUAGE CREATION PARADIGM

The previous discussion has laid out some key characteristics of teaching toward functional language use. We can now investigate their implications and special challenges for appropriate testing.

The Purposefulness of Speaking

As contrasted with much classroom language, language in the real world is purposeful. It is intended to fulfill a communicative function, such as persuading, informing, reprimanding, congratulating, or the less obvious function of establishing, maintaining, or clarifying role relationships between the communicative partners. The latter more representational function occurs through conventions that assign to a speaker certain privileges of engagement in a conversation, withholding these privileges from the other, all the while expecting both partners' behavior to mesh smoothly.

The Context-embeddedness of Language Performance

With the centrality of purpose comes the centrality of context, both linguistic and extralinguistic. We use language as a result of all our experiences in and with the world around us, the physical as well as the social world. For all the importance of the system of language for communicating, it functions only because of a context of social interaction. What behaviors have others shown in the past, what are they likely to do in the future, and what actions do we wish to take within that context?

On the linguistic side, incorporating context means using language beyond the level of simple individual sentences. Students who have only been asked to produce completely cued sentences are often helpless in conversations that span numerous turns. They have not connected the foreign language forms with a meaningful context, therefore really have never communicated in it.

In sum, if one contrasts an approach which includes the world around us with the limited engagement of students' knowledge of the world in language teaching and testing, one cannot help but conclude that teaching, as well as testing, frequently targets a use of language that does not exist anywhere else, but is totally artificial.

The Audience-relatedness of Speaking

Language in communication obtains its motivation, derives its purposes, and takes its form from the communicative partners that are involved. Closer analysis indicates that everything, from the content of a

message, to the level of its explicitness, its degree of formality or informality, the choice of words and structures, the degree of directness or indirectness, the level of involvement of the respective communicative partners, relates to the audience to whom a particular communication is directed. Again, if one contrasts the non-descriptness and nonspecificness of audience in most testing tasks with real life one finds another important area which must be attended to if communicative language testing is to become a reality.

To give a simple example: it makes all the difference in the world, even in a communicative task that is as routine and formalized as exchanging greetings, to know how old the partners are, what social status they have, how long they have known each other, what time of the day it is, how long it has been since they last saw each other, before one can judge the correctness or, perhaps better, the appropriateness of a particular language form.

Proficiency Testing - Achievement Testing - Prochievement Testing

The requirements of communicative tests spelled out thus far appear to be asking entirely too much of learners, seemingly treating them as though they already possessed complete mastery of the second language. But the real issue is the following: instead of having students invest years into learning decontextualized language forms—something that experience tells us is essentially impossible—, and then asking them to apply their knowledge in communicative settings—a transfer that, in most cases, is highly problematic—, communicative language teaching and testing assumes a context of language use right from the start. Thus, communicative pedagogy echoes first language learning by reuniting learning and use, aspects which experience, corroborated by research, tell us should never have been separated in the first place. Clearly, for the learner such use is initially possible only in

restricted settings. But within these settings the learner can strive to be as native-like as possible right from the start.

Let me further illustrate this shift by examining three terms that have become prominent in communicative testing in the American context, proficiency testing, achievement testing, and prochievement testing.

Proficiency testing refers to curriculum-independent testing that assesses the learner's ability to function in the target language, irrespective of how he or she acquired that language, or over what length of time, etc. Learners are said to possess a certain level of proficiency depending on what communicative tasks they can handle, everything from the simple interactive tasks of daily life to the highly complex formalized tasks that educated speakers must be able to handle if they wish to use the foreign language in their professional environments. Thus, proficiency testing is open-ended and a learner really does not obtain a perfect score.

By contrast, achievement testing is based on what a learner was actually taught, the materials covered within different periods of time, be that the time spent on an individual unit in a textbook or the content of a semester or an entire program of study, for instance in a high school. Ideally, test items would be taken from the syllabi on the basis of their presumed importance, as previously discussed. While there is no inherent and unalterable need to assess students' achievement in a particular unit of instruction by means of discrete-point testing, that has essentially been the kind of testing characteristic of the foreign language classroom.

This brings us to the third term, prochievement testing, an approach which attempts to combine the two. It incorporates the communicative thrust of current pedagogy and, by taking the communicative tasks students are required to perform

from a particular instructional unit, it avoids the open-endedness of proficiency testing. Prochievement testing aims to capture the interactive, purposeful audience driven, creative use of language which is the hallmark of natural language use. It does so 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.

TESTING SPEAKING ABILITY IN A COMMUNICATIVE LANGUAGE TEACHING APPROACH

We have now set the stage for a closer look at how speaking ability, in a communicative mode, might be tested in a sensible way (ACTFL, 1989).

Global Tasks

The first decision is to determine the global task a learner is to perform, the purposes for which communication is to take place. This decision has two aspects: what is the learner capable of performing, and what task is required by the communicative context. Obviously, within a pedagogical setting, the answer to the first question has greater importance. For example, a beginning learner simply is not capable of a lengthy explanation of a set of circumstances, even though the context of a debate would certainly require such an explanation. Therefore that kind of a task cannot and should not be required of the beginner. Conversely, an advanced speaker who has a much broader range of capabilities may still find it necessary to provide something like a long vocabulary list in response to the question of what grocery shopping he needs to do today: tea, fish, rice, fruit, vegetables, etc.

These seemingly simplistic facts are reflected in the hierarchy of global tasks that has been established in the so-called ACTFL oral proficiency scale which underlies much of oral proficiency testing in the United States. It shows the following progression,

where each level presupposes the ability to perform all the previous tasks:

Global Tasks

Global Tasks	
Superior	Can discuss extensively by abstracting, supporting opinions and hypothesizing
Advanced	Can describe and narrate in major time/aspect frames
Intermediate	Can maintain simple face-to-face conversations by asking and responding to simple questions
Novice	Can communicate only minimally with formulaic and rote utterances, lists and enumerations

Socially Derived Context

There exists, of course, an infinity of social contexts in which language takes place. Each culture defines its own contexts. As a consequence, an important component of communicative teaching and testing is to help students understand these culturally determined contexts.

Points to consider include:

• Who is engaged in communicating, and how many participants are there?

• What is their status, therefore who may begin the conversation, who ends it, who may question, who may not, etc. ?

• What is the setting where this communication takes place, the home, the work place, the public arena?

• How frequently does this occur?

• What background information must be considered?

One way to reduce this extensive set of considerations is to look at social setting from the following perspective: Is this an interactive context? Most of our daily use of language takes place between two people who have a certain task to perform. Or is this a non-interactive context? Presentations at professional gatherings generally end up being of that type. To some extent, the distinction between interactive and noninteractive is similar to the distinction between informal and formal settings, although not necessarily so.

What is most important about this distinction is the repercussions it has for language use. Interactive contexts are handled with context-embedded language. One does not need to explain everything since the context clarifies most things. In a family owning a poodle, the request "could you please walk the dog?" does not create the incredulous question "which dog do you mean, the German Shepherd, or the Shitsu or the Lhasa Apso?" The designation "the dog" is sufficient for referring to "the dog that we own who is a poodle, who is sitting in the next room and needs to go out." Since everyone essentially shares the same level of information, relatively reduced, perhaps even little language, is necessary. Couples who have been married for a long time are a good example of this phenomenon. Not too much talking goes on since the partners start to think the same way because of essentially the same experiences.

By contrast, non-interactive contexts tend to require decontextualized language. In a formal talk it is difficult to know the level and degree of shared information that the audience has. Some people may have one kind of information, others another. And since there is no way to clarify this through questioning and answering, the speaker must use more elaborated language to cover all the various possibilities. More

background must be provided, more explanations must be given, more options weighed, all of which requires language, therefore means much more language capability.

Let us look at how communicative language testing represents this fact in a hierarchy.

Following this hierarchy, we would ask beginning students to perform tasks that tend to be more interactive. In order for that to be possible and natural more background information must be specified which means that comparatively little language will be required to handle the task in a totally natural fashion. The task would closely mirror the situation of the beginner without being stilted and contrived.

Context

Superior	Most formal and informal settings
Advanced	Most informal and some formal settings
Intermediate	Some informal settings and a limited number of transactional situations
Novice	Highly predictable common daily settings

Also, at least two people should be engaged in this kind of testing, with the teacher preferably not being one of them, but communication taking place between students in partner and small group work. In this fashion, one of the key characteristics of an interactive context can be incorporated: the opportunity for both communicative partners to ask for details and clarification from the other, and also to ask whether they are still being understood, to request help with

their own language. These are crucially important skills to be developed by beginners. They are not signs of deficiency, but signs of competence. Therefore they are worthy of being tested.

By extension, it stands to reason that the kinds of monologic situations that we sometimes ask our students to handle should really be best reserved for advanced students. Only they possess the requisite range of grammar and vocabulary to handle these contexts effectively, and even they can only be expected to possess that ability if our instruction has provided ample opportunity for its development, something that is not always the case.

The Content

The next level of consideration as we test for communicative ability is what content our learners are likely to be able to deal with. The following progression seems reasonable:

Content

Superior	Wide range of concrete and abstract general interest topics and some special fields of interest and expertise
Advanced	Concrete and factual topics of personal and general interest
Intermediate	Topics mostly related to self and immediate environment
Novice	Common discrete aspects of daily life

Here another matter deserves clarification. It may well be true that many of our advanced learners would rather not deal with matters pertaining to daily life, preferring to handle content in their own area of professional expertise, - such as engineering, or the natural sciences. But closer examination of their language reveals the following pattern: they are likely to speak individual sentences, with little or no connectors between them. Though they will show basic awareness of the forms of the language, perhaps even possess some of the required vocabulary, success in the conversation would depend very much on a sympathetic conversational partner that shares their professional background knowledge. In other words, their ability to converse is more the result of shared previous information than it is the result of their ability to handle the language competently. This is an important distinction to make, even if we acknowledge, as we should, that any successful communication depends crucially on shared background knowledge.

To summarize, their language ability, in a general sense, is essentially at the level we have associated with the interactive speaker, only that they use professional vocabulary.

Let me emphasize that this is, of course, perfectly acceptable. Only, we should not assume that such speakers, just because they deal with professional content, can handle the language at the professional or advanced level. This confusion is, indeed, one that leads to frustration for both teachers and students.

The Text Type

These considerations lead directly into the fourth factor which must be considered, the text type. The following hierarchy regarding ability perhaps also a hierarchy of language acquisition, seems to apply:

Text Type	
Superior	Extended discourse
Advanced	Paragraphs
Intermediate	Discrete Sentences
Novice	Individual words and phrases

One reason for considering text type is because each has a high likelihood of occurring in certain situations and in conjunction with certain content. For testing this means that a learner who is essentially operating on the sentence level, finding it still quite challenging to handle all the grammatical and lexical requirements of a basic sentence in the foreign language, should not be given a task which inherently has much higher requirements with respect to text type. For example, such a student of English should probably not be asked to explain to an American educator the goals and approaches underlying the Japanese school system. Yet, these are often tasks that students get over and over again in our tests, and over and over again they lead to quite unsatisfactory results, not to mention frustration.

As a matter of fact, many teachers have even been conditioned to accept as good language something that no one outside the classroom would be willing to tolerate. Indeed, that is where the business community often rightly criticizes our profession, for not producing the kinds of speakers who use language in socially acceptable ways.

Accuracy

The final major category to be considered in communicative testing is accuracy. Its major subcomponents are pronunciation, vocabulary, grammar, sociolinguistic appropriateness, pragmatic competence, and fluency.

Traditionally, this has been the area that has received most attention, though we essentially focused on only two components, grammar and vocabulary, since pronunciation, sociolinguistic appropriateness, pragmatic competence, and fluency come into play only in actual communication.

It is noteworthy that a communicative approach to testing, particularly to testing speaking ability, gives totally different weight to these aspects of language: vocabulary and grammar are only two of five subcomponents of accuracy which, in turn, is only one of five—the others being the tasks, the content, and the text type—that one needs to consider.

Even so, accuracy is by no means irrelevant in the age of communication. That would be just as unacceptable as it was unacceptable to disregard the dynamics of communication when we focused on correct language forms. Only now, accuracy no longer leads a life of its own: it is connected to the following considerations:

• How comprehensible is the speaker with the kind of accuracy, or, in reverse, the kinds of errors being produced?

• To what extent is the burden of facilitating successful communication unfairly placed on the native speaker, who must guess, must ask, must give help, must politely cover up the long pauses and stretches of silence?

• Is the learner's inaccurate use of the language offensive to native speakers? As you know, every culture identifies very strongly with certain aspects of its language and shows different levels of tolerance for different infringements on accuracy. While some mistakes are accepted from foreigners who are beginning learners, the same mistakes would be cause for discomfort if an advanced speaker were still making them.

• Is the speech being produced so halting

Heidi Byrnes

that it becomes painful?

In sum, since accuracy exists in the service of successful communication, inaccuracies may be tolerated and tolerable as a transitional phenomenon from a learner, as long as communication is not seriously hampered.

Communicative language testing must be closely attuned to the kind of interlanguage development we discussed earlier. It must develop a careful balance between accepting the kinds of errors learners engage in, and yet not seeming to be too forgiving. It must reward learners for their successful efforts at communication, even if these efforts are not always totally accurate.

THE RELATIONSHIP OF TEACHING AND TESTING, OF TESTING AND THE CURRICULUM

So how might one go about doing this kind of testing?

After our lengthy look at the intimate relationship between the goals of language learning and teaching practice it should come as no surprise if we now extend this relationship to include testing.

As mentioned earlier, communicative testing is possible and fair only if it was preceded by communicate teaching and that, in turn, is possible only if such use of language is, in fact, the goal of language learning. Only then is it possible to establish the kind of learner outcomes which can then become the focus of teaching as well as testing.

This allows another look at the relationship between teaching and testing. Not infrequently teachers subtly subscribe to the notion that tests are most useful if they indicate a learner's ability to perform under the worst conditions. The hidden agenda seems to be to prove to the students what they cannot do, presumably to motivate

them to study even harder. In reality, such testing is more akin to a trap, to a game of chance. By contrast, prochievement testing with its communicative direction aims to allow students to prove what they can do. It accomplishes that by making an explicit connection between teaching and testing. In fact, approaches to teaching communicatively as they have been outlined can be directly transferred to approaches to testing communicative language ability. This is so since communicative teaching builds on learner involvement in a decentralized classroom. Learners will engage in various partner and small group activities whose central point is to accomplish a certain communicative task. The degree to which that task was accomplished is inherently a measure of language ability, the core of any testing.

For example, learners at a relatively early stage might handle the following task:

You would like to meet with a classmate or friend sometime over the weekend. Call up the friend on the phone. Find out when it will be convenient to get together. Make plans for what you would like to do and how and where you will meet.

This seemingly simple task requires a tremendous amount of interactive, negotiative work between the partners. In all likelihood it involves using greeting rituals over the telephone, questioning about time, about places, about things one might do, perhaps suggesting some possibilities or rejecting them as not very practicable, finally deciding on a convenient meeting time and place and, in the end, concluding the conversation with the appropriate closing ritual. With relatively little change in the parameters, such that one or the other student may not have too much time, or may not have much money to spend, the entire conversation would turn out quite differently. Thus, it could easily be used on a test: it is the same kind of task that students have practiced previously, yet, it also requires

new creativity which proves their communicative ability.

Role plays whose parameters have been specified with respect to global task, setting, content, and communicative intention can be typed up on cards ahead of time and can be used both during class as well as in testing. Over a period of time teachers can develop a repertoire of tasks which makes testing communicative ability no longer the dreaded extra task, but an extension of what has already taken place in class.

Obviously, this close connection between teaching and testing can become a tremendous motivational force for the students. The teacher's expectations will be very clear, and students will see that it is to their advantage to participate in these activities in class, even if they may have been reluctant to do so initially.

Teaching and testing will be more closely connected in another important way, namely the diagnosing of learner progress. Here the audio tape can be an invaluable aid in helping students develop their abilities. By having small tape recorders available in class two students can record their role play tasks without the teacher being present. The teacher can later collect these tapes and listen to them and provide feedback. Likewise, the students will have an opportunity to listen to themselves, perhaps even collect something like a portfolio of their progress with tapes that were produced over an extended period. In this way, they will get a much better sense of what they sound like and, ultimately, how they might improve their language use.

By informing students of the criteria for assessment as they were discussed one of the greatest oppositions to communicative testing, namely its alleged subjectivity, is also addressed. Once they have been informed about criteria for evaluation, learners are remarkably adept at assessing their own abilities and feel good about knowing

what they might do to improve, rather than just hearing that they were not good enough in some ill-defined way.

In this fashion instructional goals that may initially have been set in an abstract way turn into outcomes that are definable, for teachers and students alike. In turn, these defined and tangible outcomes can inform the setting of curricular goals : how realistic were the goals, can they be achieved in the time frame that was originally set, are they possible with some students, but not with others, is more time required, are different methods necessary, might different materials be incorporated?

Only if we perform this loop back between testing and curriculum and instruction will these new trends in testing as well as in teaching fulfill the promise toward improved language learning that they seem to hold.

REFERENCES

The ACTFL Oral Proficiency Interview Tester Training Manual. 1989. Yonkers, NY: ACTFL.

Canale, M. and M. Swain. 1980. Theoretical Bases of Communicative Approaches to Second Language Teaching and Testing. *Applied Linguistics* 1: 1-47.

Eisenstein, M. R., ed. 1989. *The Dynamic Interlanguage: Empirical Studies in Second Language Variation.* New York: Plenum Press.

Horwitz, E. and D. Young. 1991. *Language Anxiety: From Theory and Research to Classroom Implications.* Englewood Cliffs, NJ: Prentice Hall.

Krashen, S. D. and T. Terrell. 1983. *The Natural Approach: Language Acquisition in the Classroom.* Oxford: Pergamon Press.

Swain M. 1985. Communicative Competence: Some Roles of Comprehensible Input and Comprehensible Output in its Development. In S. M. Gass and C. G. Madden, eds. *Input in Second Language Acquisition*. Rowely, MA: Newbury House.

Swaffar, J. K. 1989. Curricular Issues and Language Research: The Shifting Interaction. *ADFL Bulletin* 20(3): 54-60.

Play Your VCR

Samuel Fuhrer

by Samuel Fuhrer

A French teacher has just returned to the U.S. from France with 15 videotapes full of culture and language reinforcing material that he would like to show his classes. A professor of art is regularly receiving videotapes of art exhibitions from Florence and Milano. A music teacher has a collection of videotapes of concerts of the Vienna Symphony Orchestra. A gymnastics coach in midwestern college wants to use Romanian tapes in his training program. An English teacher in Sweden wants to share his American tapes with Swedish students.

What do these educators have in common? They all share a major problem. When they summoned their audio visual center to show these tapes, they found that they could not be played on their equipment. These scenarios are repeated daily throughout the world, not only in educational situations, but in private and industrial settings as well.

The underlying reason for this problem is that television signals around the world are broadcast in different standards that are totally incompatible with one another. There are 3 major standards (PAL, SECAM, NTSC) of which there are additional substandards. The incompatabilies lie in the different line and cycle scans (525 lines-60 cycles vs. 625 lines-50 cycles). Thus, an American system will not work in France, Germany or Greece and vice versa, and tapes recorded on one system *will not play back on another.*

What are the options? There are basically 3 options.
1. Buy a digital standards converter. Such a machine which can convert a tape from one standard to another, costs from $50,000 to $300,000.
2. Send the tapes out for conversion. A good conversion (which loses some resolution) can cost from $200-$300 *per hour.*
3. Purchase a multi-standard system TV and VCR. Such a system which is fully functional in your home country also allows for play back of tapes in their original high resolution from virtually all other countries.

The above mentioned educators with the standards problem all chose option #3. They, along with hundreds of other schools have bought at least one multi-standard system from Cartridge King Co. (825 West End Avenue, New York, N.Y. 10025. U.S.A. tel: 212-749-0961).

Samuel Fuhrer, B.A., M.S., Phd is owner and president of Cartridge King Co, importer of multi-standard equipment.