

ARE COMMUNICATIVE LANGUAGE CLASSES BEING TESTED COMMUNICATIVELY?

Gary Ireland

Introduction

The Communicative approach to EFL teaching has become increasingly accepted in Japanese schools and universities in recent years. With the Ministry of Education actively promoting Communicative Language Teaching (CLT), a great deal of time and energy has been expended by both researchers and teachers in developing materials and techniques to help achieve what has been termed '*Communicative Competence*'.

However, relatively little has been published regarding development of techniques in which this communicative competence may best be tested and evaluated by the classroom teacher. In fact, an informal survey of colleagues involved in CLT at Japanese colleges and universities suggests two very clear problems. One, that there is little in the way of standardized testing and evaluation in many institutions, i.e., each individual teacher has his/her own completely independent method of evaluating their own students. And two, that many teachers are simply not testing their students in a communicative manner, and thus nullifying both the aims and principle of the course.

After a review of some of the methods currently employed and of the difficulties in designing suitable testing methods, this paper aims to offer suggestions for designing and implementing a fair and successful method for testing students involved in communicative courses.

What is communicative language teaching

At the very core of CLT philosophy is Hymes' (1972) theory of what constitutes 'communicative competence.' In accepting this concept advocates of CLT believe that the ability to use language communicatively involves both, knowledge of or competence in the language, and the capacity for implementing, or using this competence. According to Hymes, a person who acquires communicative competence acquires both knowledge and ability for language use in respect to:

1. Whether (and to what degree) something is formally possible.
2. Whether (and to what degree) something is feasible in virtue of the means of implementation available.
3. Whether (and to what degree) something is appropriate (adequate, happy, successful) in relation to the context in which it is used and evaluated.
4. Whether (and to what degree) something is in fact done, actually performed, and what its doing entails. (p. 281)

Building upon Hymes' theory of communicative competence, researchers such as Candlin (1976) and Widdowson (1978) recognized the need to address the functional and communicative potential of language rather than the mere mastery of grammatical structures. Widdowson (1978) argued that language courses should concentrate on the 'use' of language rather than the 'usage'. He defined 'usage' as, "That aspect of language which makes evident the extent to which the user demonstrates his knowledge of linguistic rules" and 'use' as, "That which makes evident the extent to which the language user demonstrates his ability to use his knowledge of linguistic rules for effective communication" (p.3).

For other writers such as Brumft and Johnson (1979) and Savignon (1983), the functional account of language use suggested by Halliday (1970, p. 145) that, "Linguistics... is concerned with the description of speech acts or texts, since only through the study of language in use are all the functions of language, and therefore all the components of meaning brought into focus" (1970. p.145), complimented Hymes's view of communicative competence. Halliday describes seven basic functions that language performs for children learning their first language:

1. The instrumental function: using language to get things.
 2. The regulatory function: using language to control the behavior of others.
 3. The interactional function: using language to create interaction with others.
 4. The personal function: using language to express personal feelings and meanings.
 5. The heuristic function: using language to learn and discover.
 6. The imaginative function: using language to create a world of the imagination.
 7. The representational function: using language to communicate information.
- (1975, p.11-17)

Learning a second language came to be similarly viewed by proponents of CLT as acquiring the linguistic means to perform different kinds of function. A great deal of

theorizing followed regarding how to implement the principles of CLT.

Though much has been published during the last twenty years with regard to method and syllabus design, for example, by Wilkins (1976), and Van Ek and Alexander (1980), no single favored method has emerged and numerous opinions exist with regard to what a communicative syllabus should look like.

However, it is fair to suggest that recent years have seen a move away from more traditionally based structural formats toward either functional, notional/functional, and task-based courses and materials. And though opinions differ regarding syllabus types, one area in which theorists do agree is that learners should spend as much time as possible involved in real communicative activities. This would typically involve learners in pair work, group work, role-playing, information exchange, and discussion activities. The target of such activities is to focus learner's attention on achieving success in the task through use of effective communication skills, and not only on the language needed for this communication.

The testing situation in practice

With communicative language courses designed to engage students in real-life communication activities, and aiming to actually enable students to use the language in real-life situations, it is reasonable to suppose that testing design should be based on evaluating these communicative skills. In other words, that the test would be predominantly constructed of speaking and listening items. However, in many cases, these skills, particularly oral skills, are being completely ignored by teachers when students are tested. Not only does it seem unfair to grade students on paper tests, which test knowledge of grammar and vocabulary, when they have focused their energies on learning to communicate more effectively, but also perhaps more importantly, it confuses students with regard to the goals of the course. Therefore, such tests completely negate the purpose of CLT courses.

Problems with designing communicative tests

There are several reasons why many teachers shy away from genuine communicative style testing. The two most commonly referred to are: (A) that the size of classes and time available make the logistics of giving speaking tests difficult to manage; and (B) that speaking tests are too subjective, and therefore it is difficult to assess students fairly. However, it seems that if improving oral/aural skills is an important aspect of, or indeed the central focus of the course, then an appropriate testing program must be put into place to evaluate those skills.

A small informal survey that I recently conducted of 20 college/university teachers in Japan, clearly demonstrates the wide variety of evaluation methodology currently being practiced. In fact, it would not be overstating the case to suggest that in many schools every single teacher has a completely different system of testing and evaluating their individual classes. Data collected from the twenty teachers surveyed (at five different institutions) shows no two methods of testing strategy were identical. This is a serious problem, which sooner or later all institutions in Japan will have to deal with. With discussion, experimentation, and an attitude of cooperation within language departments, there are two areas in which a great deal of improvement can be made.

- (A) Design of testing and evaluation programs which are clearly aimed at testing the skills practiced in the classroom.
- (B) Agreement and standardization of course targets and course testing methods.

If institutions are serious about providing good language programs that will actually teach people to speak the language, then now is the time to address both of these problems.

How are teachers actually testing?

To assess more thoroughly the trends in testing methods for communicative classes, I decided to conduct an informal survey. Twenty fellow college/university teachers at five different institutions were asked the following three questions:

- (A) How often are tests given?
- (B) What kinds of tests are given?
- (C) What percentage of overall student evaluation does testing represent?

Of the twenty teachers surveyed fifteen were native speakers (NS) and five were non-native speakers (NNS). It is not necessary here to list the entire data received, though it is extremely relevant to note that no two responses were the same. Not only did the survey show a great diversity in testing methodology, but just as importantly, it shows an enormous difference regarding frequency of testing and the value of that testing as a percentage of final evaluation. The range of testing frequency spread from once a year to every week. With regard to the percentage points allotted to testing in overall evaluation, the variation was between twenty and sixty percent.

It should also be pointed out at this stage that while many of the NS teachers are

attempting to employ some kind of communicative activity in their testing program, none of the NNS teachers used any kind of communicative items. The reason given for this was a lack of confidence in both setting up and evaluating such tests. In the case of NNS teachers, all tests were comprised of either multiple choice or short answer items. In most cases these tests were taken directly from the textbook.

Some examples of NS teacher testing methods can be seen below.

Kind of test	Frequency (Year)	% of Evaluation
Speaking (interview)	2	40
Textbook (multiple choice + listening)	4	30
Speaking (role play)	2	30
Written (multiple choice)	2	30
Written (short answer + multiple choice)	1	20
Listening	1	20
Speaking (role play)	Every	
Speaking (interview)	Week	40
Speaking (Question/answer)	1	10
Listening	2	20
Textbook (Short answer)	2	20
Speaking (Question/answer)	1	20
Speaking (Role play)	2	40

A short glance at the data above is enough to recognize the wide range of opinion in existence, or perhaps more accurately, the lack of consistency which exists regarding the testing of communicative classes. The information also demonstrates the problem which exists in most institutions—that there are usually no guidelines given to teachers regarding syllabus design and evaluation. Generally speaking, it is simply a case of unprincipled evaluation.

The survey demonstrated overwhelmingly that the tests typically used in many so-called communicative classes are basically designed to evaluate grammatical competence. This competence would reflect what Widdowson (1978) described as language ‘usage’. This kind of test is often incorporated into course books and is generally designed to measure knowledge of vocabulary, grammatical structures, and syntax. Without doubt the most commonly utilized tests of this kind are: multiple-choice tests and short answer tests.

Multiple-choice items

Multiple-choice items would typically involve selecting the appropriate word or phrase to complete unfinished sentences or questions. Generally speaking there will be a choice of four options and one option will be distinctly correct. For example:

- (1) I () never been to Egypt.
(A) had (B) have (C) haven't (D) was

Though this kind of testing presents an ideal platform for objectivity in grading, and is simple to create, it is very obviously in no way communicative. To test students involved in genuine communicative courses in this manner, has no value at all. In fact it is a contradiction of any course goals that have been set. Being able to select one word from a group placed in front of you is completely different to being able to use that word in constructing an utterance, which not only communicates an intention but, is also appropriate to the situational context. This kind of testing in itself has validity, as long as it is being used for evaluation of courses aimed at teaching grammatical competence.

As I have already mentioned, it is confusing and frustrating for students who have tried hard to adapt themselves to the requirements of a communicative class, only to find themselves being tested in the traditional manner, with no communicative component.

Short answer items

Short answer type tests are more likely to be devised by the teacher, and as the name suggests will generally require a short answer response. This could simply be a one-word response or in some cases may require several words or a short sentence. Generally speaking items may be classified as passage items and will usually be based on a reading passage. They are in some cases also based on spoken or visual stimulus. Though this kind of testing is still very popular, particularly in the design of college/university entrance tests, it actually falls down as a classroom format, both as an objectively gradable test and as a communicative test.

1. It is not objective.

For example, If we are to evaluate a single short answer on a written test, do we accept poor spelling but not poor grammar? In a spoken response, are we listening for correct pronunciation? Or is the proper expression, badly pronounced, acceptable?

2. It is not communicative.

Even when this test takes on an interview style format, where the teacher asks questions of the student, it is still not a communicative test. This type of test requires a learner to display his/her competence by answering questions that have no particular context or sequence and do not represent the successful filling of any information gap.

It is also worth noting at this point that just because a test has a speaking component, it does not automatically make it a communicative test.

Considerations in test design

As with designing a communicative language syllabus (Ireland, 2000), several factors must also be taken into account when designing a suitable testing program.

Apart from the aforementioned problems of class size and objectivity of tests, other factors would certainly include, sociocultural environment, student's background and experience, and individual personality.

I will attempt to deal with each of these issues in order.

Student numbers and time constraints

Without doubt, one of the biggest difficulties facing language teachers in Japanese institutions is dealing with the sheer size of classes. Sadly, in many cases, where administrators and department leaders still live in the era of Grammar Translation methodology, the size of so-called 'speaking/communication' classes will be completely at odds with the reality of teaching a communicative class. Personally, I have taught 'speaking/communication' classes ranging from ten to sixty learners, and was once actually offered a class for a hundred and twenty! It is certainly a serious problem.

It is therefore immediately obvious that simply having one testing (and of course, teaching) program will simply not suffice for all classes. Both the type and especially frequency of tests, need to be adapted to each individual course. In larger classes it may be advisable to regularly monitor and evaluate each student's classroom work, rather than trying to give end of semester tests. This can be done by testing several students each week as they do the tasks set in class. In addition, it may be preferable to include more listening items than speaking items.

Another idea is to give end of semester tests for very large classes over two class periods, in order to allow each individual sufficient time. When it is necessary to do this, the

students either waiting their turn for a speaking test or similarly, students testing in the second period, can be given listening items, or failing that, written communicative items, to utilize the time. The key to dealing with class size problems is the ability to adapt to each individual class, and not to try to implement one standard program.

Objectivity of grading

Detractors of communicative style testing argue that such tests are too subjective to be judged fairly and accurately. Whereas the traditional pen and paper type tests consisting of either multiple choice or short answer type items are objective, in that they can be graded purely on a right or wrong basis, the communicative test requires interpretation as to whether and to what extent the task given has been carried out successfully. Naturally, this requires a good deal more thought and preparation at both the planning and evaluating stage. Something perhaps that many teachers do not have the time to commit themselves to. However, communicative tests can be designed to allow good deal of objectivity. In considering how we may create and evaluate our tests, two factors are crucial:

- (A) That the test has a clearly defined and realistic target or goal in which the learner may demonstrate his/her communicative competence.
- (B) That there is a set of well defined set of criteria in place for evaluating how successful the task has been carried.

By following these guidelines and also evaluating and recording grades at the time of the (speaking) test, a good deal of the so-called guesswork or subjectivity can easily be eradicated from the process. In many cases, teachers who are reluctant to give communicative tests, have simply not thought enough about criteria for evaluation. Naturally, experience with such tests leads to more confidence in deciding to what level a task has been successful.

In considering whether a degree of subjectivity is acceptable in testing, we must remember that communication itself is subjective. If, for example, you make the same statement to fifty different people, it is fair to assume that how each person interprets the meaning of this utterance, and thus their response to it, will vary somewhat.

This will of course depend upon the interpreter's own knowledge of the language and situation, his/her sociocultural background, and perhaps personality and character. In this respect the classroom teacher who has prepared students for certain tasks is obviously in the best position to judge if these tasks have been carried out effectively.

Sociocultural considerations

As all teachers working in Japanese institutions are well aware, the average Japanese student is unlikely to want to initiate conversation, and in particular to ask or answer questions voluntarily. In a society where individuality and expressing one's own opinion is looked upon as rude and arrogant, and all decisions/ideas are formulated by group consent, it is obvious that coaxing people to communicate in English in the manner of the West is a very real challenge. Indeed, informing each new class that they will be given mainly oral tests never fails to produce a stunned reaction very worth recording on film! In fact, to suggest that the very color drains from a few of the now fraught looking faces, would not be an exaggeration.

However, it must be remembered that if the course goal is to improve communicative competence, and to engage learners in pair and group activities to achieve this, then the testing program should, in all fairness to those learners, be designed to determine the degree of that competence. In my own experience, the initial panic is soon overcome once students have been given a clear outline of the course goals and an explanation of why this style of testing is necessary. In fact, in most cases, students actually come to enjoy the tests and relish the challenge.

It is argued that communicative testing is too difficult for Japanese students because of the sociocultural difference in the way that they would usually communicate with others. However, it should be remembered that students are being prepared to go out into the outside world and to communicate in English, in a manner, which is consistent with, and acceptable to native speakers. Language and culture simply cannot be separated from each other (Ireland 2000).

If learners are being exposed to communicative, information gap type activities, then it is reasonable to expect them to be tested by completing the same kind of activities. There are already many people in Japan, with relatively high EIKEN and TOEIC qualifications, who cannot hold even a very basic every day conversation! These people are clear evidence that knowledge of a language, i.e., grammar, and vocabulary, does not necessarily facilitate ability to use the language for communication.

Learner experience

Many of the students sitting in first year college/university classes will have had little or even no experience of CLT. They may have been exposed to question and answer sessions by a NS teacher in high school, but will have had little real chance to try to

communicate more than simple agreement or disagreement. Considering the importance placed on entrance examinations and the nature of those examinations, this is not surprising—even though the Ministry of Education continues to promote the use of CLT in Japanese schools.

The one positive aspect to be taken from the current ability of freshman students is that they may be considered not as beginners in English, but as false beginners. As such, though they may not at first be capable of completing communicative tasks, they do have a base of grammar and vocabulary already in place, to be able to do so.

CLT cannot simply replace the teaching of structure and vocabulary (Ireland 2000), and therefore, communicative competence cannot exist independently of such knowledge. Where required, it will often be necessary to reinforce this knowledge of vocabulary, grammar and syntax in order to undertake communicative tasks. Bearing this in mind, it is easy to recognize that false beginners are perfectly suited to being thrown into the deep end of the communicative pond and expected to swim.

Personality

In oral communicative tests, in which self-confidence in expressing oneself, ability to gesture and the use of good intonation all play a role in achieving success, it is quite obvious that an extrovert will have the advantage over a more introverted personality. Though this factor worries some teachers, it is quite fair and normal that students who are most capable of getting their message across, and achieving the set communicative goal, will score higher on this kind of test.

However, it should be remembered that tests should be designed to measure both accuracy and fluency. In addition, the oral test may be only one part of overall testing, and that this testing itself will only make up a part of the mark for overall evaluation.

Thus, while it may be an advantage to be outgoing and ready to take risks and make mistakes, this alone will not guarantee success, when overall evaluation is taken into account. However, from my own experience, I would suggest that the more outgoing students generally do enjoy communicative tests, and often wish other teachers would adopt them.

This then leads us to consider what we mean by a communicative test.

What constitutes a communicative test?

An outline of what actually constitutes a communicative test is given by Harrison (1983). For Harrison there are three main considerations when preparing tests:

1. A communicative test should assess language used for a purpose beyond itself. An oral interview for example can be used to assess how well the learner can manipulate language in response to stereotyped questions, but a communicative version demands a response to circumstances. This rules out tests in which the learner is asked to display his language competence for no other reason than to have it assessed.
2. A communicative test should depend on the bridging of an information gap. It has to propose a language-using purpose which can be fulfilled by the communicative skill so far acquired by the learner. He must *need* to know or to tell—and his interlocutor must be in a similar necessity of explaining or finding out. For example, one kind of role-play consists of reading out the alternative parts in a script; another could be planning a holiday on the basis of travel brochures. The former has no intrinsic value as a communication: the latter necessarily involves expression of information by one part to the exchange and understanding of it by the other. This approach rules out exchanges in which the questioner has no interest in the answer (e.g. “And how did you come to school this morning (not that I care)?”).
3. Communicative test should represent an encounter. The situation at the end of it should be different from what it was at the beginning, and this means that there has to be some sequence within the test. It may require a simple transition from not knowing to knowing (e.g. understanding a text, whether written or spoken). A more complex transition would be from data given in one form to a conclusion reached in another, by way of several intermediary stages. The sequence in this case can be either putting together information contributed through several participants, as in jigsaw exercises and games, or building on several kinds of information supplied to one person, as in taking a phone message, looking something up in a reference book and leaving a written note. In this kind of exchange some personal commitment has to be made by the learner to the outcome of the communication, and the participants’ real or assumed characters and attitudes will have rubbed off on one another to some extent, however small.

It is obvious that when designing tests to reflect the notional/functional/task-based nature of CLT, the focus of those tests must reflect the style of learning which has been practiced within the syllabus. In communicative classes the majority of classroom time should be spent working in pairs and groups, and will concentrate on role-plays and information gap activities. Naturally, the focus is on listening and speaking for communica-

tion, and the amount of reading and writing involved will be small by comparison. Taking this into account then, is it not clear that testing should also be based on oral/aural tasks, rather than the reading/writing type of items mentioned above?

In addition, the tasks included in tests must reflect a communicative challenge, which has been practiced in classroom time in a period previous to testing. Also considering that most classroom activities are designed to be carried out by pairs of students, it is appropriate where possible to test students in pairs conversing with each other. In such cases, both speakers should have equal responsibility in seeing that the communicative task assigned is successfully completed.

Designing tests

Naturally, in designing tests, all of the factors discussed above must be taken into consideration, and testing programs must be adjusted to suit each individual class or institution. Perhaps the most significant point to consider being the size of group. Given the ideal circumstances of a relatively small class (up to fourteen students), it should be quite possible to create a testing program based upon communicative tasks.

Given this situation, my own preference is to give two, mid-term, and two, end-of-term tests. In the case of beginner/intermediate level students these would normally be directed conversation type tests, where the students will be informed in advance of the task they are to complete. In almost every case, students will conduct the tests in pairs.

Examples of tests include:

Arranging to go out with a friend (A calls B and invites B to go out).

Discussing plans for the weekend/or summer/winter vacation.

Discussing a weekend or vacation in the past tense.

Discussing opinions (topic given in advance).

Asking for/giving advice (problems decided by students on testing day).

Talking about experiences.

Ordering dinner in a restaurant.

Starting a conversation with a stranger (situation given in advance).

There are two crucial considerations to ensure that this kind of testing is successful. (1) As already mentioned, the tests should be based upon tasks that have been completed in the classroom. Students should be familiar with both the function they are carrying out, and the

relevant vocabulary/structures that will be necessary to carry out their part of the task.

For example, in ‘Arranging to go out with a friend’, students will have practiced (with several partners) during class time. They will have been given an introduction/revision of language used for inviting/suggesting, and will have had the chance to build their knowledge of vocabulary and useful language by completing the task several times in realistic conversation practice. This can be done, for example, by having students move around the class and fill up their free time for the weekend with appointments.

During classroom practice, students will have been taught the language necessary to make the conversation as realistic as possible, i.e., to actually say what they normally would say—“I don’t feel like a movie” or “I can’t stand Chinese food”—and not simply to use the easiest language available to complete the task.

This use of language should of course be taken into account when the test performance is evaluated.

(2) Students must have been moved around the classroom regularly, and have had chance to work with a large group (preferably the whole class) of different individuals. For successful CLT to take place (especially in Japan), it is imperative that students are exposed to variation, and not allowed to become part of small groups (Ireland 2000). It is simply no use at all to allow anyone to continually sit with and work with a best friend only, and not have to talk to anyone else in class.

If the practice of moving places/partners regularly is implemented from the beginning, students will accept it as quite normal to do pair testing with any of the individuals in their group.

The importance of this partner rotation cannot be stressed too strongly. It serves several purposes. Talking with different people helps to build confidence in one’s ability to communicate. It also allows students to interact with a wide range of personalities rather than only their own favorite group. This in turn exposes students to a variety of experiences, opinions and ideas. This all guarantees that students will not know what to expect each week from their partner.

In classes of the size mentioned above, it is not difficult to organize a change of partners on a weekly basis. Naturally, with larger classes, the time and organization required dictates that a less regular change is necessary. Perhaps bi-monthly, or every three weeks as a worse case scenario.

Another key to success in this kind of test is to select testing partners randomly at testing time. Sadly, many teachers tend to allow students to stay in the same seats with the same partner year round, and even to test together as a pair. This is not only poor

technique for CLT, but also nullifies the communicative aim of such a test. Naturally, given the opportunity to do so, students will simply prepare and rehearse a dialogue to produce on the day. Unfortunately, instead of a spontaneous conversation, where both partners need to understand the other's meaning and get across their own intentions, what takes place is simply a monotone, memorized and meaningless skit.

In addition to the four main tests suggested above, students can also be given regular evaluation tests at the end of a class period. This can be in the form of shorter, directed conversation type tests or perhaps short presentation type tests, for example, outlining a plan for a proposed journey, or telling a short story of a personal experience. Naturally, the task will be based upon work done in the classroom. In larger classes, where holding four main tests is not practical because of time restraints, this type of testing can be very helpful in generating sufficient grades for evaluation. In addition, this relatively brief testing can involve different students in each session.

Alternatively, this mini test format is an opportunity to introduce listening tests into the program. The ability to understand the other person's intentions is just as crucial as the ability to get across one's own message, and therefore, it is appropriate to test this skill.

Although a case could be argued for multiple-choice items in listening tests, particularly to provide a source of objective results, once again there are much better methods of evaluating aural comprehension. Short answer type items can be utilized, where students are required to listen for information can also be effective for giving objective grades.

However, where the task concerned involves listening for information, it is more appropriate for students to have to create full responses to the aural information input. This could be either, by giving full answers to questions, or listening to dialogue and completing segments of a conversation. The focus of the task being again to both understand the context of the aural message, and to respond in a manner that would be appropriate.

Personally however, where possible, I prefer to use the listening format to test the ability to understand the feeling of the speaker rather than just to listen for information. This means creating tests in which, listening for stress and intonation to determine a speaker's true meaning is the target. Examples of this kind of activity can be found in Jack Richards et al, *Listen for it* (1987). In this text, these exercises are called 'listening tactics'. An example of one exercise is listening to people accepting invitations and deciding by their intonation whether the speakers are genuinely interested or not in the proposal. This kind of activity/test requires not only an understanding of the language used, but also an understanding of the speaker's feeling towards what is being said.

Naturally, like all other testing items, these exercises should be carried out in classroom time if they are to be used successfully in a test.

Criteria for grading communicative courses

Though a good deal has been published regarding selecting criteria for grading communicative tests, particularly by J.B. Carroll (1980), and J.A. Upshore (1971), very little of what has been written is by or for the ordinary classroom teacher. Most of the research done to date concerns proposals of criteria for either English for special purposes (ESP) courses, for example, for translators or interpreters, or else has been designed for implementation by institutions or testing bodies. Because of this, most of the test criteria suggested are far too complicated and time consuming to be practical for use in ordinary once-a-week communicative classes. When the teacher has to listen to (and watch) two testees simultaneously and grade their performance at the same time, having ten different categories to consider is impossible. The test has to be designed to be as straightforward as possible.

The generally accepted theory is that the focus of testing criteria should be designed to evaluate 'fluency' and 'accuracy'. In addition to this measure of oral input, I also prefer to allot a grade for the aural aspect of the task. This can simply be titled, listening comprehension, and is a mark that reflects how well the testee has understood his/her partner's utterances.

For the verbal content then, i.e., the 'fluency and accuracy', it is sufficient to divide the grading areas into the following four segments:

Fluency of speech;

To what extent the conversation is carried out naturally and without excessive delays or revision of utterances.

Naturalness of discourse:

To what extent the intonation and rhythm is natural and realistic, and not monotone and/or memorized.

Accuracy of language used:

To what extent, and how accurately did the student utilize the language learned and practiced during class time.

Successful completion of task:

To what extent the testee was able to complete his/her role in the communicative task assigned.

Each of these areas and 'listening comprehension' are given a grade from a scale of 0 to 10. Therefore, the final grade for each test will be from a maximum of 50 points, which can then be easily transformed into a percentage figure. Where four main tests (ideal case) have been given, or a combination of four or more main and shorter tests, a reliable average testing evaluation can be achieved. It is quite possible for any of the students to have one poor test, but over a series of tests the level of their ability should be clear to the teacher. Again, it should be remembered that having designed, set up, and monitored the classroom activities undertaken, the classroom teacher him/herself is in the very best position to judge how well an individual has learned to deal with those tasks.

Overall evaluation

It is clear from the example survey results above that a variety of opinion also exists as to what percentage of the overall evaluation criteria should be allocated to test results. In fact, the range of points allocated extended from 30 to 60 percent. In every case however, the additional evaluation criteria consisted of a combination of points allotted or subtracted for attendance, classroom attitude and effort, and homework/special projects. Again, the percentages points allocated to each factor varied widely.

Though it is not within the scope of this paper to analyze overall evaluation methodology, it is worth noting that this is another area in which a good deal of discussion is needed to reach a more standardized approach.

Generally speaking, I prefer to allocate 50 percent to testing, 20 percent to attendance, 20 percent to effort and attitude and 10 percent to homework/special projects. Since without regular attendance, and sufficient classroom effort, one cannot prepare for this kind of testing, the system tends to balance itself very well, and of the thousands of students evaluated in this manner so far, none has ever questioned a grade.

Conclusion

Though some institutions have begun to ask teachers to use a standard text, or in some cases to work towards common goals, generally speaking, most institution's communicative English language programs could be improved enormously by simply clarifying and standardizing both the aims of the course, and the manner in which the course is to be evaluated. If successful CLT is a genuine concern to any institution, it is no longer acceptable to simply allow each individual to continue interpreting it's meaning in their own way.

Initially, department leaders need to sit down with teachers and reach a clear understand-

ing on both what the goals of the courses are to be, and how best to realize those goals. Through discussion and co-operation, it should be possible to implement far more standardized and useful programs.

In addition to course development, there is an urgent need for both more research and more experimentation at the classroom level, as to how these communicative courses can be most efficiently and fairly tested and evaluated. Testing programs must be designed and employed which reflect the concept of the approach.

If CLT is to be successful in Japanese institutions, it's time to start teaching what we preach, and testing what we teach.

References

- Brumfit, C.J.(1980). From defining to designing: communicative specifications versus communicative methodology in foreign language teaching. In K.Muller, (ed.), *The Foreign Language Syllabus and Communicative Approaches to Teaching: Proceedings of a European-American Seminar*. Special issue of *Studies in Second Language Acquisition*, 3/1, 1-9.
- Brumfit, C.J. & Johnson, K. (eds.),(1979). *The communicative approach to language teaching*. Oxford: Oxford University Press.
- Candlin, C.N.(1976). Communicative language teaching and the debt to pragmatics. In C. Rameh (ed.), *Georgetown University roundtable 1976*. Washington, D.C.: Georgetown University Press.
- Halliday, M.A.K (1970). Language structure and language function. In J. Lyons (ed.), *New Horizons in Linguistics* (pp. 140-165). Harmondsworth: Penguin.
- Halliday, M.A.K (1970). *Explorations in the functions of language*. London: Edward Arnold.
- Hymes, D.(1972). On communicative competence. In J.B. Pride & K. Johnson (eds.), *Sociolinguistics*. (pp.269-293). Harmondsworth: Penguin.
- Harrison, A.(1983). 'Communicative testing: jam tomorrow?' in Hughes and Porter (eds.), *Current Developments in Language testing* 1983: (pp 77-85). London: Academic Press Inc.
- Hughes, A. & Porter, D.(1983) *Current developments in language testing*. London: Academic Press Inc.
- Ireland, G.V.(2000). Eclecticism: Balancing Communicative Language Teaching With the Japanese College/University Classroom. In Bulletin of the Department of English Language and Literature Bunkyo Women's College. 33 (pp.65-83).
- Richards, J.C., Gordon, D., and Harper, A.(1987) *Listen for it*. Hong Kong: Oxford University Press.
- Savignon, S.(1972). *Communicative competence: theory and classroom practice*. Reading, Mass.: Addison-Wesley.
- Widdowson, H.G.(1978). *Teaching language as communication*. Oxford: Oxford University Press.
- Wilkins, D.A.(1976). *Notional syllabuses*. Oxford: Oxford University Press.

Bibliography

- Bachman, L. F.(1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Brown, H.D.(1994). *Principles of language learning and teaching*. New Jersey: Prentice Hall.

- Carroll, B. J.(1980). *Testing communicative performance*. London: Pergamon Institute of English.
- Richards, J.C., & Rodgers, T.S.(1986) *Approaches and methods in language teaching*. Cambridge: Cambridge University Press.
- Underhill, N.(1987). *Testing spoken language*. Cambridge: Cambridge University Press.
- Upshore, J.A.(1971). Productive Communication Testing: a Progress Report. In G. Perrin and J.L. Trim (eds.). *Applications in Linguistics*. Cambridge: Cambridge University Press. (pp.435-442).
- Valette, R. M.(1971). Evaluation of learning in a second language. In S. Bloom, J. Thomas, Hastings and G.F. Madaus (eds.). *Handbook on Formative and Summative Evaluation of Student Learning*. (1971). New York: McGraw-Hill.
- Van Ek, J.A., & Alexander L.G.(1975). *Threshold level English*. Oxford: Pergamon.
- Wadden, P. *A handbook for teaching English at Japanese colleges and universities*. New York: Oxford University Press.
- Widdowson, H.G.(1979). The communicative approach and its applications. In H.G. Widdowson (ed.). *Explorations in Applied Linguistics*. Oxford: Oxford University Press.
- Weir, C.J.(1988). *Communicative language testing*. Exeter: University of Exeter.
- Yalden, J.(1983). *The communicative syllabus: evolution, design and implementation*. Oxford: Pergamon.