DOCUMENT RESUME

ED 226 567 FL 013 198

AUTHOR Gallo, Gary A.
TITLE Developing Models of Communicative Competence: Conceptual, Statistical, and Methodological Considerations.
PUB DATE 82
NOTE 17p.; Revised version of paper presented at the Annual Convention of Teachers of English to Speakers of Other Languages (16th, Honolulu, HI, May 1-6, 1982).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Communicative Competence (Languages); Language Proficiency; Language Tests; *Models; Research Methodology; Statistical Analysis

ABSTRACT

The development of an empirically based model of communicative competence is discussed in terms of conceptual, statistical, and methodological considerations. A distinction is made between descriptive and working models of communicative competence. Working models attempt to show how components of communicative competence are interrelated psychologically to form a set of independent factors and to determine whether aspects of language are psychologically distinct and can be learned separately. In reviewing the research on this question, several methodological and statistical issues must be addressed. Among these issues are interpretation of the Pearson product moment correlation coefficient, subjects' language background, patterns of language proficiency, and criterion versus norm referenced language measures. Examples are provided of the ways in which these issues can influence the interpretation of patterns of test results. Research on the components of communicative competence are reviewed and the findings reinterpreted in light of the methodological issues. (Author/RW)

Developing Models of Communicative Competence: Conceptual, Statistical, and Methodological Considerations

Gary A. Cziko

University of Illinois

Running head: Developing Models of Communicative Competence

Abstract

The purpose of this paper is to discuss a number of conceptual, statistical, and methodological considerations which have an important bearing on the development of an empirically-based model of communicative competence (CC). First, the notion of components of CC is examined and a distinction is made between descriptive and working models of CC. Second, statistical and methodological considerations are discussed, including issues related to interpreting correlation coefficients, norm-referenced vs. criterion-referenced measurement of language skills, and the language background of language learners included in empirical studies of CC. Third, some important studies on the components of CC are reviewed and their findings reinterpreted in light of these considerations. Finally, topics and guidelines for further needed empirical research in this area are offered.

Developing Models of Communicative Competence: Conceptual, Statistical, and Methodological Considerations

Over the last several years, a good deal of thought and research has been concerned with the problem of discovering the linguistic and extra-linguistic abilities which make up communicative competence (CC), i.e., the ability to communicate effectively and appropriately in a given language in various situations. While there now appears to be general agreement among applied linguists, sociolinguists, and psycholinguists that far more than grammatical knowledge or competence is required for CC, there has been a continuing debate between proponents of a model of CC which puts emphasis on one general factor of CC (e.g., Oller, 1976; Oller & Hinofotis, 1980) and those who maintain that a number of distinct components must be included in a valid model of CC (e.g., Bachman & Palmer, 1982; Canale & Swain, 1980; Cummins, 1981). However, apparently lost in this debate are a number of important conceptual, statistical, and methodological considerations.

The purpose of this paper is to bring to light a number of these considerations and to discuss their implications for interpreting both theoretical and empirical work in this area and for planning additionally needed empirical research. First, the notion of components of CC is examined and a distinction is made between descriptive and working models of CC. Second, a number of statistical and methodological considerations are discussed which have a crucial bearing on the interpretation of empirical research on working models of CC. Third, several important studies on the components of CC are reviewed and their findings reinterpreted in light of these considerations. Finally, topics and guidelines for further needed empirical research in this area are offered.

Components and Factors of Communicative Competence

Central to the development of models of CC is the notion of components of CC. Essentially, most theoretical and empirical work in this area has been concerned with defining what the components of CC are and describing how these components are interrelated. However, there are two very different ways of going about this task. Depending on one's perspective and purpose, one may wish to consider the components of CC to be representative of all the knowledge and skills necessary to communicate effectively and appropriately in a given language. In contrast, one may be instead interested in determining how many psychologically independent or orthogonal types of knowledge or skills make up CC. Since the statistical technique of factor analysis has commonly been used to determine how many statistically uncorrelated components account for the variance of a group of subjects on a large battery of psychological measures, factors will be used in this paper to refer to components or combinations of components of CC which are statistically independent (uncorrelated) for a given group of language learners. On the other hand, components will refer to any types of knowledge or skills which appear to be involved in CC in a given setting, whether or not these sources of knowledge are psychologically or statistically independent of each other (i.e., form separate factors). Failure to keep this distinction clear between components and factors of CC appears to be responsible for much of the controversy and confusion which currently exists in this area.

The former perspective of describing and defining all of the components of knowledge and skills that a person needs to communicate effectively and appropriately in a given language can be described as developing a descriptive model of CC. The most widely known current descriptive model of CC appears to be that developed by Canale & Swain (1980) and Canale (in press). In this descriptive model, CC is divided into four principal types of competence, viz., grammatical, sociolinguistic, discourse, and strategic competence. For a descriptive model of CC, what is of primary concern is the inclusion of all types of knowledge that are necessary to be successful in communicating via language and making sure that the model's components are organized into a logical and intuitively appealing structure. Thus, while Canale and Swain (1980) first included discourse competence as a part of sociolinguistic competence, Canale (in press) now considers discourse competence to be distinct from sociolinguistic competence and worthy of separate billing. Presumably this change was not based on findings from empirical research and does not necessarily suggest that sociolinguistic and discourse competence are separate or independent in any statistical or psychological sense, but simply points out that for descriptive models of CC, describing components and their interrelationships is primarily a problem of definition.

In contrast, working models of CC attempt to show how components of CC are interrelated psychologically to form a set of independent factors. Of principal interest from this perspective is to determine whether the fine divisions which linguists and sociolinguists use to describe language (e.g., phonological, morphological, syntactic, discourse, and sociolinguistic rules) are psychologically distinct so that a language learner can learn them separately or learn more of one than of another. This, then, is an empirical question, answerable only by carefully designed and implemented research. While a fair amount of this type of research has been done, (see, e.g., Oller, 1976; Oller & Hinofotis, 1980; Bachman & Palmer, 1982), there is still considerable debate over the findings and their interpretation (see Oller, 1981). Before this research is reviewed, however, there is a need to describe the basic research methodology used in this type of research and the problems inherent in interpreting the results of such research. In fact, there appear to be some very basic statistical and methodological considerations which have been ignored in designing these studies and interpreting their results.

Statistical and Methodological Considerations in Developing Empirically Based Working Models of Communicative Competence

Interpreting the Pearson Product-Moment Correlation Coefficient

It appears that all of the research designed to investigate the factors of CC and their interrelations have been correlational in nature. To illustrate, let us take a simple example of a study that might be done to investigate the relationship between English-as-a-second-language (ESL) reading and speaking skills involving 100 foreign students studying at an American university. The primary question to be examined is whether reading and speaking skills are related (i.e., a student who is strong in reading is also strong in speaking while a student who is weak one skill is also weak in the other) or independent (i.e., knowing that a student is strong or weak in one skill does not allow one to accurately predict ability in the other skill). To answer this question, a researcher would typically

locate (or construct) tests of English reading comprehension and speaking, administer them to the group, and calculate the Pearson product-moment correlation coefficient to describe the degree of relationship of scores on the two tests. If the correlation is high (.7 or above), a typical researcher would likely conclude that these two skills of English CC are highly related, while if the correlation coefficient is low (.4 or below) he is likely to conclude that the two skills are relatively independent of each other. However, there are a number of important considerations which should be kept in mind before coming to these conclusions.

First of all, while a low correlation coefficient may well indicate that there is little if any relationship between the two skills under study, the opposite interpretation is also possible. Since the correlation coefficient tells us how much of the variance on one test is predicted or accounted for by variance on the other test, the amount of interindividual variation on each test also influences the value of the correlation coefficient. Therefore, if all of the students in our group score uniformly high on both tests, the value of the correlation coefficient may be quite low in spite of the fact that no student scored high on one test and low on the other. Conversely, while a high correlation coefficient may well indicate that there is a strong relation between reading comprehension and speaking skills, the opposite interpretation may also make sense. This could be the case if the students ranged from educated native-speaker proficiency down to an intermediate level of proficiency in reading comprehension while they ranged from an intermediate level down to virtually no proficiency in speaking English. The fact that each student is notably weaker in speaking ability would suggest that the two skills are not highly related in the strongest sense of the word and that one general language proficiency factor would not adequately explain the performance of these students.

Yet, even with most students performing notably worse on the measure of speaking ability, there may still be a high correlation coefficient between reading and speaking skills since the coefficient is based on relative performance of each student compared to his group and not on any absolute or criterion-referenced level of language proficiency. This brings into question the exclusive use of norm-referenced tests for this type of research and suggests that the use of criterion-referenced language tests which yield scores which are meaningful in terms of known levels of language proficiency may be more appropriate (see Cziko, 1981 and in press).

Language Background

Another factor which must be considered in the interpretation of the results of language tests is the type of English-language background shared by the subjects under study. English-language background refers to the type of contact the subjects have had with the English language and the amount of opportunity they have had for acquiring the various skills and components of English CC. This background can vary in three basic ways which are portrayed by the patterns shown in Table 1. In these patterns, points toward the top of each band represent individuals with a high degree of exposure to a given language skill while points toward the bottom of each band represent individuals with a low degree of exposure to a given language skill. Lines connect the same individual's position in the two language skills portrayed (i.e., reading and speaking). While the language background patterns of Table 1 include only seven out of many more possible patterns, only four individuals per group, and exposure to only two language skills, extensions to other patterns, larger groups, and more skills could be easily made.

The first way in which a group's language background pattern may vary is in the degree of homogeneity of the group with respect to the amount and type of exposure they have had to each of the ESL skills under consideration. For example, the only exposure that a group of foreign students may have had to English before coming to study in the U.S. could have been that provided by a series of formal courses taught at the secondary school and university undergraduate levels. If all of these students had taken the same or very similar courses, then they would be very homogeneous in terms of their exposure to English and it could be said that they have low within-skill variance in their exposure to those ESL skills they had had the opportunity to acquire. This low degree of within-group variation is shown in the patterns of Examples 3, 4, 5, and 7 of Table 1 where there is relatively little vertical spread among the four points for each language skill. If, on the other hand, the group consisted of students who were very heterogeneous with respect to their exposure to one or more ESL skills (e.g., some had gone abroad for a year to study at a British or American university while others remained home and took only a few English courses) they would then be characterized by high within-skill variance in one or more ESL skills. Groups with a high degree of within-skill variation in exposure to ESL language skills are shown in the patterns of Examples 1, 2, and 6 of Table 1 where there is a large amount of vertical spread among the four points for the two language skills. To keep Table 1 relatively simple, not shown are groups characterized by a high degree of within-skill variation in exposure to one skill and low within-skill variation in exposure to another skill.

Insert Table 1 about here

Second, a student may have had the type of exposure to the English language which fostered the development of all of the components which make up English CC. Such a background could be provided by a foreign school or university which has a strong English program, requires students to take courses in their area of study using English as the medium of instruction and discussion, requires students to study from English texts, and provides contact with native speakers of English via student exchanges or programs of study abroad. In contrast, the only English courses another student may have had could have been concerned solely with the development of skills for reading technical, scientific English. Hence, a student may have a very even or uneven exposure to the various components which are believed to be important to CC. A person with even exposure to these components will be referred to as having low between-skill variance in ESL background. Groups made up of such individuals are shown in the patterns for Examples 1, 3, and 5 of Table 1. Low between-skill variance in exposure to the skills of reading and speaking is shown in these patterns by the low slope of the lines connecting the points (representing the same individuals) in the two language skills. In contrast, groups comprising individuals characterized by high between-skill variance in exposure to the two language skills are portrayed by points connected by lines having high slopes in Examples 1, 4, 6, and 7 of Table 1.

Finally, the relationships of exposure to ESL across different skills may vary considerably from group to group from strong positive correlations through weak or close to zero correlations to strong negative correla-

tions. For example, a group of students may have the option of taking English courses for one, two, three, or four years with all courses designed to develop both reading and speaking skills. In this case there would be a strong, positive correlation between exposure to English reading and speaking skills since those students who study English longer will get more exposure to both skills. The resulting language background pattern would be that described in Examples 1, 3, 6 or 7, depending on the amount of within-skill and between-skill variance in exposure to the two skills. Conversely, if the students had available a large variety of equally popular English courses, each one with a different skill focus, then there would be a low correlation between amount of exposure to different English language skills resulting in language background patterns similar to those of Examples 2, 4, or 5 of Table 1. Although negative correlations are also possible and indeed likely for certain groups of subjects, they are not considered in Table 1 for reasons of economy. Table 1, therefore, represents just seven of a much larger number of possible language background patterns which can result from combinations of the above three aspects of within- and between-skill variance in language background and the correlation of exposure between language skills.

Now that these three aspects of language background have been described, let us examine how one of these may influence the interpretation of the data in our example. If our 100 foreign students have a high positive correlation in amount of exposure to opportunity to acquire English reading and speaking skills (as in Examples 1 or 6), we should not be surprised to find that there is a fairly high correlation between their reading and speaking test scores. This would not necessarily indicate, however, that the two skills are dependent on each other or are simply two manifestations of a unitary underlying language ability. Those students who have devoted more time to studying English would be expected to do better in both reading and speaking than those who have had less time for English study if their English study had given them exposure to both skills. If, on the other hand, we find a low correlation between reading and speaking test scores for this same group, this would seem to indicate that reading and speaking are in a very real sense separate language skills. Such a finding would suggest that we could not predict English reading ability based on observed ability to speak English and that teaching one skill would not necessarily influence the other.

Let us now consider a group for which there is no correlation in exposure between the skills being measured, as in Example 2, 4, or 5 of Table 1. If for this group we found a low correlation between language skills it would seem to make most sense to attribute this to the language background of the group and not to any psychologically real independence of the language skills under consideration. If, however, we find a high correlation between reading and speaking skills, this would appear to provide impressive evidence for the interdependence of these two language skills, an interdependence which as not an artifact of language background.

Although we have only considered the correlation in exposure between language skills in interpreting the results of language test score patterns, both within- and between-skill variance in exposure to language skills also should be considered in making these interpretations. However, a discussion of these factors will be more meaningful after the concepts of the next section are introduced. Nevertheless, the general point being made

here can be expressed quite simply. That is, the pattern of results of language tests administered to group of second-language learners can only be meaningfully interpreted in light of the language background of the group. Instead of focussing solely on the pattern of test results, patterns of test results should be compared with the language background pattern of the group. If this is done, then we may well find that what is often taken as evidence for either a one-factor or multi-factor working model of CC may instead be simply an indication that the pattern of language proficiency one acquires is related to the type and amount of exposure to the language one has had. If, however, it is found that the pattern of test results is not congruent with the language background pattern of a group, this would then appear to be important evidence for the psychological reality of the model which best explains the test result pattern.

Patterns of Language Proficiency

We now need to consider how language skill proficiency patterns can vary across groups of language learners. As for the language background patterns described above, patterns of language proficiency can vary in at least three basic ways; examples of seven possible patterns of language proficiency involving the two skills of reading and speaking are given in the column headed "True Skill Pattern" of Table 2. The patterns of this column are analogous to those of Table 1 except that the vertical dimension now represents language proficiency level. Points at the top of each band represent individuals with language proficiency equivalent to that of a well-educated native speaker of the language while points at the very bottom of each band represent individuals with no proficiency in the language skill under consideration.

Insert Table 2 about here

The first way in which language proficiency patterns can differ across groups of individuals is in the amount of within-skill variance in proficiency in given language skills. A homogeneous group with respect to language proficiency would have little within-skill variance in proficiency in language skills. Patterns characterizing such homogeneous groups are shown in Examples 3, 4, 5, and 7 of Table 2. In contrast, more heterogeneous groups characterized by high within-skill variance in language proficiency are shown in Examples 1, 2, and 6 of this table.

A second way in which language proficiency patterns can differ is in the amount and direction of correlation between language skills. This correlation, usually quantified by use of the Pearson product-moment correlation coefficient, can vary from close to negative unity (indicating a strong inverse linear relationship) through zero (indicating no linear relationship) to close to positive unity (indicating a strong positive linear relationship). For the sake of brevity, only patterns characterized by strong positive linear relationships (Examples 1, 3, 5, and 7) or linear relationships close to zero (Examples 2, 4, and 5) between the two language skills are included in Table 2, although it is possible that an inverse relationship (negative correlation) could exist between proficiency in two language skills.

Finally, we need to consider the amount of between-skill variance in proficiency between language skills. In Table 2, this is indicated by the amount of slope of lines connecting individuals' proficiency levels across the two skills of reading and speaking. Thus, in Examples 1, 3, and 5

there is little or no between-skill variance while for Examples 2, 4, 6, and 7 the between-skill variance in proficiency between reading and speaking is considerable. If measures were used which reflected individuals' true level of proficiency on language skills from zero, representing no proficiency, to some arbitrary fixed maximum, representing the proficiency level of an educated native speaker of the language, then true between-skill variation in language proficiency between two skills could be mathematically defined as the sum of squares of the differences between each individual's pairs of scores on the two skills divided by the number of individuals in the group.

This concept of true between-skill variance in proficiency in language skills would appear to be of critical importance in determining the number of factors underlying the CC of a group of language learners. For instance, if our group of learners showed high between-skill variance between skills in English reading and speaking (as in Examples 2, 4, 6, and 7 of Table 2) it would appear that a one-factor interpretation of the pattern is inadequate since it fails to take into account the fact that native-speaker level proficiency in reading is not uniformly associated with native-speaker proficiency in speaking. Conversely, low within-skill variance between language skills (as in Examples 1, 3, and 5 of Table 2) would appear to support a one-factor interpretation since the proficiency level of each individual would be approximately the same in each of the language skills.

Criterion- and Norm-Referenced Language Measures

While this discussion of true skill language patterns has obvious importance for developing models of CC, it must be realized that completely valid information concerning the true skill proficiency patterns of a group of language learners is never possible to obtain. Instead we must build our theories and models using the information provided by measures of language proficiency, all of which introduce error into our assessments of language proficiency levels. There are, however, some important differences between criterion-referenced (CR) and norm-referenced (NR) measures of language proficiency with respect to the type of information they can provide concerning the true skill proficiency pattern of a group of language learners.

For the purpose of this discussion, a CR measure of language proficiency will be considered to be one which yields scores which are directly interpretable in terms of a language performance standard (see Cziko, 1981, p. 32). Examples of such measures are the oral interview and reading tests originally developed by the Foreign Service Institute (FSI) of the Department of State (see Jones, 1979) and now administered by the Interagency Language Roundtable (ILR) as well as the ESL dictation test and copytest developed by Cziko (in press). These CR language measures contrast with NR measures of language proficiency which yield scores which are not interpretable in themselves, but rather must be converted to a standard score or percentile rank which in effect compares an individual's score with the scores of all other examinees (usually other second language learners) on the measure. Examples of such NR measures are the Test of English as a Foreign Language (TOEFL) developed by the Educational Testing Service and typical applications of integrative testing procedures such as the cloze and dictation procedures when the tests are not calibrated with respect to a specific criterion group of language users.

This distinction between CR and NR language measures is of critical importance whenever one attempts to determine the true skill proficiency pattern of a group of individuals based on their performance on language measures. Since CR measures are designed to yield scores which are interpretable with respect to a specified language performance standard (e.g., educated native speakers of the language), reliable and valid CR measures of language skills could be expected to yield a pattern of test scores which are useful approximations of the underlying true skill proficiency pattern of the group being studied. Thus, if our group of language learners all scored between 4 and 5+ on the ILR test of French reading but between 0 and 1+ on the ILR French oral interview, this group's true skill pattern for reading and speaking French would likely be similar to the true skill pattern shown in either Example 4 or 7 of Table 2 (the exact pattern depending on whether or not a high correlation was observed between the two measures). However, since for NR tests a group's mean score as well as its within-test variation are typically considered arbitrary, the scores resulting from such measures are usually standardized so that the mean and variation is made to equal certain prespecified values. The most common standardization procedure is the z transformation which linearly transforms raw test scores so that their mean is equal to zero and their standard deviation and variance are equal to unity. It should be noted that this standardization procedure is involved whenever the Pearson correlation coefficient between two measures is calculated, as is done for all statistical procedures based on the general linear model such as regression analysis and factor analysis (both exploratory and confirmatory).

This standardization of the mean and variance of a set of test scores has two important effects on the type of NR test result patterns that one can obtain. First, NR tests and correlational analyses standardize the within-test variance so that regardless of the true within-skill variance the amount of adjusted within-test variance is the same for each language test. Consequently, as shown in Table 2, the differences between true skill patterns of Examples 1 and 3 as well as between 2 and 4 are not revealed in the corresponding NR test results patterns. Second, since means are standardized as well, any true skill patterns characterized by differing levels of overall proficiency on two or more skills (i.e., high between-skill variance) are adjusted on the NR test results pattern so that these differing levels of proficiency are made equal (i.e., between-skill variance is reduced). This standardization of group means together with the standardization of within-test variance described above has the effect of eliminating the differences between the true skill patterns of Examples 4 and 5 as well as between Examples 6 and 7 of Table 2. It should be particularly noted that if the correlation between two language skills is high, NR test results pattern will always show low between-test variance in spite of the possibility that the true between-skill variance may in fact be quite high (as in Examples 6 and 7 of Table 2); while if the correlation between two language skills is close to zero, then the NR test results pattern will invariably show high between-test variance in spite of the possibility that the true between-skill variance may in fact be quite low (as in Example 5 of Table 2). Thus, it becomes apparent that while CR measures of language skills have the potential to accurately reflect the true skill pattern of a group of language learners with respect to a number of language skills, NR measures as well as any statistical techniques based on correlation coefficients are seriously limited in this respect. The implications of these differences between CR and NR measures of language proficiency for

the development of working models of CC will be considered along with other factors in the next section.

Interpreting Patterns of Test Results: Seven Illustrative Examples

We have now seen that there are five important considerations which influence the interpretation of the results of studies which attempt to describe and validate working models of CC. These include (a) the Pearson product-moment correlation coefficient between two tests and its interpretation, (b) the language background pattern of the examinees, (c) the amount of within-skill variance in language proficiency, (d) the amount of between-skill variance in language proficiency, and (e) the type of tests used to obtain measures of language proficiency (NR or CR). We will now begin a systematic examination of how combinations of these considerations can affect our interpretation of the results of studies which attempt to describe and validate working models of CC.

To do this we will again use the simple situation of a group of ESL learners who have been administered only two tests--tests of English speaking and reading comprehension. We will also continue to use Table 2 which shows seven examples of combinations of within-skill variance, between-skill variance, and between-skill correlations as well as simplified graphic illustrations of true language skill patterns, NR test results patterns, and both NR and CR interpretations for each of the seven examples. (The importance of language background characteristics will be considered later.) Note that the essential difference between the true skill patterns and their respective NR test results patterns is that the latter are based on standard scores which in effect equate the means and variances of the two tests considered in each example when these are not already equal in the true skill pattern. The CR test results pattern for each example is not shown in Table 2 since if we were in fact able to use CR language measures according to the definition stated above, the resultant test patterns would be very similar to the depicted true skill patterns, differing only by measurement error.

For the first four examples, there is no difference between the NR and CR interpretation of the test results with respect to the number of independent factors accounting for the test results. This is because in these four examples, high between test correlation coefficients are accompanied by low between-test variance (Examples 1 & 3) and low between-test correlation coefficients are accompanied by high between-test variance (Examples 2 & 4). Therefore, regardless of which of these statistics is used as a criterion for the number of independent factors underlying ESL reading and speaking skills (the correlation coefficient for the NR interpretation and the between-test variance for the CR interpretation), the interpretation would be the same.

For the fifth example, however, the NR and CR interpretations conflict. The NR interpretation of two factors would be based on the low correlation coefficient between the two tests while the CR interpretation of one factor would be based on the low between-test variance. From an NR point of view, it makes sense to say that since relative performance on one test does not predict relative performance on the other, there must be two independent skills responsible for the NR test results pattern. However, this interpretation does not consider that in absolute terms there was very little between-skill variance to begin with. This is taken into account by the CR interpretation which would conclude that since all the examinees

seem to be at the same general level on both language skills, one general proficiency factor likely underlies both skills.

Finally, while the NR interpretation of the sixth and seventh examples is straightforward due to the high correlation coefficient between the two tests, the CR interpretation appears problematic. In both of these cases we have high between-skill variance since all examinees performed better on the reading test than on the speaking test in absolute terms. However, each examinee performed at the same relative level on the two tests. In these cases, both one- and two-factor interpretations appear to make some sense but neither interpretation alone appears to do justice to the results pattern that would be given by CR tests. Since NR interpretations of test result patterns are undoubtedly more common (since little else is reported in the literature), most readers would likely be impressed with the high between-test correlation and give any interpretation other than a single-factor one little consideration. Nevertheless, the high between-test variance must be taken into account if CR language tests were used which accurately reflect the true skill pattern of proficiency in the two skills. Not to do so would mean neglecting the important findings that for the group under examination, native-speaker level proficiency in reading ESL is not associated with native-speaker level proficiency in speaking ESL. Obviously, something more is involved in speaking -- something more than is accounted for by simply saying that both skills are related to a single language proficiency factor. Perhaps it could be said that for the sixth and seventh examples some sort of one-and-a-half factor interpretation makes the most sense!

Now, how do language background factors affect these seven illustrative examples? As stated earlier, a one-factor NR interpretation of language skills would seem to be less than convincing if based on data from a group having a high correlation between exposure to the ESL skills being assessed, while a multi-factor NR interpretation of language skills would seem ill-founded if based solely on data from a group with a low between-skill correlation in exposure. With respect to Table 1, all one-factor NR interpretations of test result patterns would be most convincing if based on data from groups similar to those shown in Examples 2, 4, or 5 while all NR multi-factor interpretations would be best founded on data provided by groups similar to Examples 1, 3, 6 or 7. If, however, we make use of CR language measures and wish to make CR interpretations of test result patterns based primarily on the amount of between-skill variance in test results, then a one-factor interpretation would appear to be best founded if based on groups which had high between-skill variance in exposure to language skills (as in Examples 2, 4, 6, and 7 of Table 1) while a multi-factor interpretation would appear to be most convincing if based on groups having low between-skill variance in exposure to language skills (as in Examples 1, 3, and 5 of Table 1). It is also important to keep in mind that data and interpretations obtained from one group cannot be generalized to other groups unless a strong case can be made that the groups share a common language background pattern.

Another thought which this discussion brings to mind is that it may be senseless to speak of a model of second-language CC without reference to a particular group of second language learners. Surely, for many foreign students who enter American colleges and universities, different aspects and modalities of ESL skills may be highly correlated lending credence to a one-factor model of CC (see Oller, 1976; Oller & Hinofotis, 1980). However, this certainly does not necessarily imply that there are not many individuals who are much stronger in some skills or aspects of

skills than in others. Indeed, no one would argue that there are not many individuals throughout the world who can speak a second language and yet not read a word of it or others who can read materials in a second language with little difficulty yet can hardly utter a word of the language for the purpose of oral communication. It would appear, consequently, that researchers would be wiser to investigate the relationship of language background factors to second language proficiency rather than attempt to derive general models of CC based on groups of learners whose second language backgrounds are either unknown or peculiar to the group under study. Additional points relevant to this discussion will be made in the next section which reviews some of the most influential research in the areas of CC and language testing in light of the considerations raised in this paper.

A Review of Selected Studies

In the preceding sections, all illustrative examples were kept as simple as possible so as to not confuse the basic concepts and considerations being discussed. However, to show that the considerations expressed in this paper do have an important bearing on actual empirical research done in the area of CC and language testing, some of the most influential research in these areas will now be reviewed.

The first report of empirical research designed to systematically collect evidence for a general language proficiency factor is that of Oller (1976). In this research, Oller administered a large battery of ESL tests to foreign students enrolled at the University of California, Los Angeles. In spite of the fact that the tests administered had been designed to measure different aspects of ESL proficiency (e.g., vocabulary, grammar, reading, dictation, and cloze ability), it was found that the between-test correlation coefficients between all possible pairs of the tests were fairly high (from .54 to .79) and, consequently, the individual tests all loaded highly on a single factor of a principal components analysis (a form of factor analysis). In order to review Oller's conclusion that these findings constitute evidence for a single general language proficiency factor, however, we need to examine information concerning the types of tests used in this research and the language background of the participating subjects.

The language tests used in this research are described in some detail in Oller (1972) and Oller and Streiff (1975). It is clear from these descriptions that the tests used (all subparts of the UCLA ESL Proficiency Examination) were designed to be norm-referenced measures of ESL proficiency. That is, they were constructed to produce inter-individual variance but not designed to yield scores which are meaningful with respect to any criterion level of language proficiency. With respect to the ESL background of the subjects, however, little information is given. Had the majority been from the same country it might be reasonable to assume that they shared a homogeneous language background (i.e., low within-skill variance in exposure to ESL skills) that was either even (fairly equal exposure to the various components of communicative competence in English, i.e., low between-skill variance) or uneven (unequal exposure to these components, i.e., high between-skill variance). On the other hand, had they been from a variety of countries, a heterogeneous language background would be more likely (i.e., high within-skill variance on exposure to ESL skills). Unfortunately, this language background information is not given in spite of the fact that it is difficult to interpret the results of these studies without it.

Relating this research to the seven examples of skill and test patterns described earlier, it appears that Oller obtained data which resemble the NR test results pattern depicted in Examples 1, 3, 6, and 7 of Table 2. However since CR language tests were not used, there is no way of knowing which of the four true skill patterns shown in those four examples comes closest to depicting the language skills of these students. If the true skill pattern given in the first or third example underlies the test data, then the one-factor interpretation given by Oller may well be the most reasonable. If, however, the true skill pattern of the sixth or seventh patterns is closer to the true nature of things, then the one-factor interpretation of the results would be considerably less convincing. (Note that although these studies involved five language tests and not just two as depicted in Table 2, the use of this table as an aid in examining these studies is nevertheless appropriate since all the statistical analyses reported by Oller are based on the relationships of results of all possible pairs of tests.) So it appears that while the NR interpretation of these results is clearly a one-factor one, the use of CR language tests and the analysis of the between-skill variance might have led to a quite different interpretation (the so-called one-and-a-half factor interpretation mentioned above).

There may be even more reason to suspect the validity of the one-factor interpretation of these test data if we consider the probable English-language background of these students. If they had shared a heterogeneous, even English language background (i.e., high within-skill variance and low between-skill variance in exposure to ESL skills), then the relatively high between-test correlation coefficients obtained in these studies could very well have been due to the fact that those students who had more exposure to English had the opportunity to develop proficiency in all five of the skills measured by the UCLA tests while those who had less exposure to English naturally did less well on all five of the language measures. Without more information on the type and amount of these students' exposure to English, however, it is impossible to know if this is a likely explanation for the reported findings. It should be clear, however, that there are sufficient reasons to be wary of a one-factor interpretation of these results.

In a later publication, Oller and Hinofotis (1980) examined the correlations of a number of ESL tests administered to foreign university students for evidence concerning the divisibility of components of ESL skills. The data reported in this study are particularly interesting in that they give some indication of the possible effect of language background on language test result patterns and because CR tests were used in part of the study. With respect to language background, it was found that ESL test data collected from a group of 159 Iranians studying at the University of Tehran gave clearer support for one general factor of language proficiency (from a NR perspective) than did data collected from a (presumably) more heterogeneous group of foreign students at Southern Illinois University. If the Iranian students in the Tehran study had taken the same general type of ESL courses but for varying lengths of time, then there likely would have been stronger intercorrelations among exposure to these skills than for the foreign students in the U.S. That the test results of these two groups appear to resemble what we would expect from their

language background is an indication that language background factors can have important influence on test result patterns and must be taken into consideration when test result patterns are interpreted to support either a one- or multi-factor view of CC.

In addition, Oller and Hinofotis made use of the FSI oral interview procedure to obtain CR measures of English accent, grammar, vocabulary, fluency, and comprehension for students in the Illinois study. The use of these CR language measures would have made it possible to get a clearer picture of the true skill pattern of the students than that provided by NR measures. However, none of the statistical analyses used to analyze these data was sensitive to the absolute level of proficiency indicated by the FSI scores. Instead, all analyses were based on statistical procedures which reduced between-skill variance accompanied by high intercorrelations between test scores. It is therefore possible that a re-analysis of these data using a CR definition of between-test variance based on the original raw scores of the FSI subparts might reveal that the data were not as easily accountable for by a one-factor model of CC.

The most recent empirical work on the factors of CC is that of Bachman and Palmer (1982). The work of Bachman and Palmer distinguishes itself from other research done in the area for its use of different statistical techniques (causal modeling and confirmatory factor analysis instead of the more commonly used exploratory factor analytic techniques) and a very heterogeneous sample of ESL learners (representing ages from 17 to 67 years, 18 different native-language backgrounds, U.S. residence from a few months to over 10 years, and formal as well as informal ESL learning contexts). Although the statistical analyses used by Bachman and Palmer do appear to have considerable advantages over exploratory factor analytic techniques (see Oller, 1981), these techniques nevertheless share the same limitation mentioned earlier of reducing between-skill variance which is accompanied by high intercorrelations between test scores. Thus, these techniques do not make it possible to detect any of the true-skill patterns shown in Examples 4, 6, or 7 in Table 2. This is the case despite the fact that a number of the language measures used in this research (the three FSI ratings and possibly the three self-rating scales) yielded CR information.

It is therefore difficult to know how to interpret Bachman and Palmer's analyses which suggest the existence of a general factor accompanied by two additional specific trait factors of CC. The fact that Bachman and Palmer used a sample of subjects which appeared considerably more heterogeneous in language background than those included in the previous work of Oller and his associates also raises the possibility that the inadequacy of one general factor to account for these results may have been at least partly due to language background factors. If one general factor could not explain the pattern of exposure of these language learners to various ESL skills (as is the case for Examples 2 and 4 of Table 1), then is should not be surprising to find that one general factor fails to adequately account for their pattern of language test results.

Concluding Remarks

The arguments of this paper lead directly to the conclusion that in spite of the amount of thinking, research, and writing that has been done on CC, we still have but a very fuzzy picture of an empirically based

working model of the factors composing CC. It should be clear that if additional empirical research is to further our understanding of the factors comprising CC, it must involve the use of CR language measures designed to reveal the true skill patterns of language learners together with appropriate statistical analyses and extensive language background information must be collected and considered in the interpretation of the results provided by these measures.

References

Bachman, L. F., & Palmer, A. S. The contruct validation of some components of communicative proficiency. TESOL Quarterly, In press.

Canale, M. A theoretical framework for communicative competence. In J. Richards & R. Schmidt (Eds.), Communicative competence. Rowley, Mass.: Newbury House, In press.

Canale, M., & Swain, M. Theoretical bases of communicative approaches to second language teaching and testing. Applied Linguistics, 1980, 1, 1-47.

Cummins, J. The role of primary language development in promoting educational success for language minority students. Paper prepared for the California State Department Education Compendium on Bilingual-Bicultural Education, 1981.

Cziko, G. A. Psychometric and edumetric approaches to language testing: Implications and applications. Applied Linguistics, 1981, 2 (1), 28-44.

Cziko, G. A. Improving the psychometric, criterion-referenced, and practical qualities of integrative language tests. TESOL Quarterly, In press.

Jones, R. L. The oral interview of the Foreign Service Institute. In B. Spolsky (Ed.) Some major tests. Arlington, Va.: Center for Applied Linguistics, 1979.

Oller, J. W., Jr. Scoring methods and difficulty levels for cloze tests of ESL proficiency. Modern Language Journal, 1972, 56, 151-158.

Oller, J. W., Jr. Evidence for a general language proficiency factor: An expectancy grammar. Die Neueren Sprachen, 1976, 2, 165-174.

Oller, J. W., Jr. Language testing research (1979-1980) In R. L. Jones & G. R. Tucker (Eds.), Annual review of applied linguistics: 1980, Rowley, Mass.: Newburg House, 1981.

Oller, J. W., Jr., & Hinofotis, F. B. Two mutually exclusive hypotheses about second language ability: Indivisible or partially divisible competence. In J. W. Oller, Jr. & K. Perkins (Eds.), Research in language testing. Rowley, Mass.: Newburg House, 1980.

Oller, J. W., Jr. & Streiff, V. Dictation: A test of grammar based expectancies. In R. L. Jones & B. Spolsky (Eds.) Testing language proficiency. Arlington, Va.: Center for Applied Linguistics, 1975.

Footnote

This is a revised version of a paper presented at the annual convention of Teachers of English to Speakers of Other Languages, Honolulu, May 1982. Requests for reprints and other correspondence should be addressed to the author at the University of Illinois, Educational Psychology, 1310 S. Sixth Street, Champaign, IL 61820.