25 ● Translation Accuracy: Using Multidimensional Scaling

GEORGE A. BARNETT ● MARK T. PALMER ●
HANA NOOR AL-DEEN

*State University of New York — Buffalo ● University of Wisconsin —
Madison ● State University of New York — Buffalo*

A CCURATE language translation is a vital issue in bilingual and multilingual societies. National development and social progress rely to a great extent on the mutual understanding and coopera- tion of linguistic groups. The problems resulting from the lack or distortion of information transfer across cultural boundaries affect the vital areas of sci- ence, industry, government, and education. The level of technology and production achieved is directly dependent on the quality of available infor- mation. Ethnic minorities and other culture-bound groups, when deprived of the opportunity to develop or adopt to technological innovations, are handicapped in modern competitive economies. Policy decisions in busi- ness and government can also be severely distorted by poor communica- tion through inaccurate translation, which can limit political participation and thus the process of national integration.

Finally, intercultural understanding as a whole suffers from poor-quality translation because it places groups and individuals into positions of high un- certainty. If information is inaccurate or uninterpretable, the potential for mu- tual understanding and trust diminishes and the potential for conflict in-

Correspondence and requests for reprints: George A. Barnett, Department of Communica- tion, State University of New York, Buffalo, NY 14260.

creases. This chapter tests a method to ensure quality translation that has the potential to facilitate intercultural communication and mutual understanding.

BACKGROUND[1]

Communication can be thought of as a continuous process in which information is shared between individuals or groups. The convergence theory of communication describes this process as being directed toward the goal of mutual understanding (Kincaid, 1975; Rogers & Kincaid, 1981; Barnett & Kincaid, 1983; Palmer, 1981; Palmer & Barnett, 1982; Kincaid, Yum, Woelfel, & Barnett, 1983). Participants continue to exchange information as long as there are perceived differences in their understanding of each other or of the content of their communication.

In a single-culture context, this can be represented as a unique, functional system of concepts or meanings (a system of symbols) that facilitates the social interaction of the members. When communication is limited to a particular topic, the reference frames are also limited for the participants.

> When we communicate with others we are endeavoring to present our concepts in the domain of our conversation and to learn from them about the separate relations among concepts in the same domain. (Woelfel & Fink, 1980, p. 37)

In an intercultural or bilingual context, mutual understanding is the result of the same convergence process, the increased similarity between the semantic/conceptual frameworks of the participants from different cultures or linguistic groups (Barnett, 1974; Palmer & Barnett, 1984).

Translation is a vital means of communication between language-bound cultures. From the convergence perspective, mutual understanding may be represented in translation as the equivalence of two texts (one being the translation of the other). If the texts are substituted for the individuals participating in the convergence process, equivalence is the goal of the translation process. This results when the semantic/conceptual domains relevant to the content of the texts are the same. Accuracy or clarity in translation is then a function of bringing the domains into equivalence.

Therefore, the fundamental issue lies in how we judge equivalence. What are the indicators of the degree of similarity or difference in the conceptual/semantic frameworks of the participants and how can an observer know when the cyclical process has reached its goal of equivalence or mutual understanding?

Ideally, the bilingual translator stands at the meeting of the two languages and their cultures (MacNamara, 1970; Nida, 1975; Barnett, 1977a,

1977b). However, it is not always the case that a bilingual is trained in the particular subject matter dealt with in the material to be translated. Also, there is no guarantee that one translator is equally competent in both languages. As a result, equivalence of technical documents should not be determined by the judgments of bilinguals. Some empirical measure of the cognitive/semantic frameworks involved in translated texts must be used. This measure must take a behavioral approach using responses of a sample of the intended users of the text to the actual text itself. These individuals should be viewed as judges or observers of the text rather than as subjects involved in an experiment. Their responses should be taken to be an evaluation of the translation.

Back-translation has thus far shown the best results in bilingual/expert judged situations (Ervin & Bowers, 1952; Ervin & Osgood, 1954; Brislin, 1972; Brislin & Sinaiko, 1973). Typically, in this method a document is translated back into its original language and both original-language versions are compared by bilinguals acting in the place of single-language users. Using the same cyclical process, but modifying it to include objective measures, would eliminate the variance accounted for by bilingual judges.

The most effective methods of representing the relationship between concepts (meaning) in semantic domains as they are embedded in language have been developed through spatial modeling using either the semantic differential scale or multidimensional scaling (Osgood, Suci, & Tannenbaum, 1957; Osgood, May, & Miron, 1975; Barnett, 1976). These techniques present the relationship, through the mathematical manipulation of data, along comparative dimensions and across linguistic boundaries. The result is a mathematically defined space describing the semantic/conceptual frameworks of the language groups in relation to the text. This is simply a way of comparing concepts that users of the texts consider important. Because there are several problems that make the semantic differential scale less than ideal as a measure of meaning and therefore as a measure of equivalence, metric multidimensional scaling is preferred to determine the degree of semantic/conceptual equivalence (Barnett, 1976; Woelfel & Fink, 1980).

Proposed Methods

The most versatile measurement method available for determining the quality of translation is the Galileo system of metric multidimensional scaling (MMDS). It meets the qualifications specified above and has been shown reliable and theoretically valid in several linguistic and cross-cultural studies (Barnett, 1974, 1977a, 1977b, 1979, 1980; Barnett & McPhail, 1980; Wigand & Barnett, 1976; Barnett, Wigand, Harrison, Woelfel, & Cohen, 1981; Kincaid et al., 1983).

The process of measuring semantic/conceptual frameworks begins with determination of the key concepts in the document to be translated. This is done through in-depth interviews with representative samples from the population of intended users in the original language after they have read the document. For the translation of a technical document, these individuals would be professionally equivalent to the intended users in the target language.

The second step is to measure precisely the relationship among the key concepts to generate the semantic/conceptual space of the users. This is done using the original text, a tentative translation, and back-translation as stimuli. Measurement is done using the method of pair-comparisons (Thurstone, 1927). Subjects estimate the dissimilarity between the relevant concepts against a criterion standard. Typically, the questions are worded in the following way: If X and Y are U units apart, how far apart are concepts A and B?

This format allows respondents to report any positive number, rather than limiting their responses by forcing them to choose a point on a fixed scale, as with Likert and Osgood's semantic differential scales (Barnett et al., 1981). This increases the variability of responses, their precision, and the scale's ability to describe change over time. Further, it allows for direct comparison among the symbols, requiring none of the assumptions of the semantic differential.

The completion of the data collection operations results in a concepts $\times$ concepts $\times$ subjects matrix (S) that is square and symmetrical. Each cell in the matrix (s_{ij}) represents the distance between concepts i and j. To determine the linguistic characteristics of a group, S is averaged. Any entry s_{ij} is the mean distance between concepts i and j, as seen by the average member of the group (Barnett, 1977a, 1977b).

Matrix S may be converted to a multidimensional space. Mathematically, the process is analogous to converting a matrix of intercity distances to a Cartesian coordinate system where latitude and longitude are the reference axes and the cities' locations on these dimensions are given. From the coordinates a graphic representation such as a map may be drawn. In that special case, an n $\times$ n matrix of cities may be reduced to a two-dimensional configuration with little loss of information. However, in semantics, this may not be the case. Semantic space is often multidimensional, and when comparing the semantic/cognitive frameworks across languages for equivalence, all the dimensions used to differentiate the symbols should be taken into account (Barnett & Woelfel, 1979).

Change or difference in language may be examined by repeating the pair-comparison phase and transforming the data for each language. Translated equivalents of the key concepts are used for the language into which the text is being translated. Comparisons of measurements in both languages make it possible to examine the differences between the texts and

how individual concepts differ across languages. In this way one can determine where the translation is in error and how to manipulate the concepts to modify the translated document to achieve equivalence.

This assumes that the semantic/cognitive space generated from readers of the original document represents a fixed criterion against which the translation is evaluated. If this assumption is made, then the translator may engage in a cybernetic process of reducing the differences between the spaces (error in translation) through an iterative process of testing and adjusting the translated document (Palmer & Barnett, 1982, 1984). Alternatively, the original document may be adjusted to make it more amenable to the translation process. This is known as "pretranslation" (Nida, 1966). This dynamic interplay between the original document and its translation, which would result in their congruence, is more consistent with the tenets of convergence theory. It is appropriate for many nonscientific materials, especially, governmental documents in multilingual societies.

AN EMPIRICAL EXAMPLE

The goal of this research is to evaluate the utility of metric multidimensional scaling as a procedure to determine the accuracy of translation of written materials in natural language. Once the translation's accuracy and the locations of discrepancies from equivalence are known, the researcher may intervene to ensure that both documents have a common meaning. This section will demonstrate these procedures through a translation of a document from English to Arabic. No specific hypotheses will be presented or tested because this research is evaluative and therefore descriptive in nature.

Methods

Step 1:
Selection of the Document for Translation

The first step in the process is the selection of a document for translation. The following selection criteria were employed:

(1) The document must be precise and technical to control for variance in interpretation. That is, the relations among the concepts in the text must be clear.
(2) There should be a limited number of key concepts. They should be clear, standing out from the surrounding text, which should describe the relationships among these symbols. There should be a high level of agreement among the subjects as to which are the key concepts in the text.
(3) While the document should be technical, it must at the same time be translatable. Symbols for the concepts in the text must have counterparts in the lan-

guage into which it is to be translated. Since English has become the *lingua franca* of science and commerce, many documents cannot be used because the English symbols for many technical concepts have been borrowed directly.

(4) The document should not be mathematical or contain graphics. The text should avoid universal symbols; it should not present information in a manner that need not be translated.

(5) The document should not be too difficult for the subjects to comprehend. It should present a complete, unambiguous set of relations in a relatively short space—two or three pages. This will prevent subject mortality and ensure subject motivation to complete the testing procedures.

(6) The document should not be in the subjects' area of academic experience. In this way, subject equivalence across language can be assumed. Simply, the subjects can be matched using knowledge of the subject matter as the criterion.

The selected document was a portion of Ball-Rokeach and De Fleur's "A Dependency Model of Mass Media Effects" (1982, pp. 157-158). It met all the criteria mentioned above. The only limitation on the subject selection was that subjects be unfamiliar with mass communication.

Step 2:
Selection of the Samples

The samples for this study included a group of subjects for which Arabic was the first language and a group of English speakers. The 50 Arabic speakers were enrolled at the State University of New York at Buffalo. The majority were Palestinian. Others came from throughout the Middle East and North Africa. The majority were engineering majors, but some were studying physics, linguistics, education, management, sociology, pharmacy, computer science, geography, and chemistry. There were no communication majors among the Arabic sample.

The English speakers were drawn from introductory communication classes and were excused from the research if they had taken a mass communication course or were not native English speakers. The total English-speaking sample was 208. These subjects performed three tasks: (1) They identified the key concepts in the text; (2) they estimated the relations among the selected key concepts in the original text; (3) they estimated the relations among the key concepts after reading the back-translation.

Step 3:
Pretesting to Determine the Key Concepts

After the text and samples were chosen, a portion of the English sample (N = 31) read the text and identified the "key words" or "key concepts," as well as the topic sentence or the theme of the text. The key concepts and topic sentence were elicited because it cannot be expected that single words

alone in any language will convey the meaning of a text or translate directly into single words in another language. Since equivalence is measured through the relationships among concepts, word-for-word equivalence need not be expected.

A content analysis of the responses revealed that the most frequently mentioned words or phrases were as follows:

word or phrase	frequency
(1) audience	18
(2) mass media	24
(3) society	19
(4) dependency	29
(5) information	20
(6) social conflict and change	10
(7) needs	9
(8) fantasy and escape	7
(9) cognition	5
(10) feelings	6
(11) behavior	8

These concepts formed the basis of a Galileo questionnaire. There was variation in consensus regarding how "key" each symbol was. While "dependency" was selected by 29 of 31 respondents (94 percent), "cognition" was chosen by only 5 (16 percent). This lack of agreement on the document's central concepts may prove problematic.

Step 4:
Translation of the Text

Portions of the essay were translated from English to Arabic through standard back-translation procedures. One of the authors translated the document into Arabic. She is an Iraqi studying for a doctorate in communication. Her course work has focused upon the mass media. She has one master's degree in behavioral science and another in English. Her Arabic translation was translated back into English by an Egyptian doctoral student majoring in educational administration, with course work in communication.

There were substantial differences between the original English text and the back-translation due to variations in dialect between Egypt and Iraq and differences in expertise on the topic. Among the discrepancies were semantic variations involving the words "public," "people," and "society." These differences were resolved in discussions until there was agreement on the sample translation. The back-translation was edited to make it grammatical.

The translation process resulted in three versions of the text: the original English version, an Arabic translation, and a back-translation in English. This allowed for the comparison between the original text and both the translation and the back-translation. Comparisons between the original and the back-

translation have certain advantages. First, often translations are performed in one society and then sent to another. In that case, sufficiently large samples in the second language may be unavailable. Comparisons using the back-translation allows the researcher to evaluate the translations for discrepancies without leaving the society where the original was produced.

The second advantage of comparing the back-translation to the original is that cultural differences between the two language groups do not enter into the evaluation of the translation. All subjects come from the same culture. Thus the only source of variation across the two conditions is the text, either the original or its back-translation.

However, the goal of the translation process, following convergence theory, is for the two different language groups to reach a common understanding of the meaning of the text, despite cultural differences. For this reason it is essential to focus on the responses of the actual users of the translation in their own language. Thus it is necessary to compare the original text and the translation directly.

Step 5:
Perform Galileo Analysis in Original Language

A Galileo instrument was created by randomly ordering the concepts from the pretest and then forming all possible pairs of words or phrases. The criterion pair was "needs and dependency are 10 units apart." This instrument was given to two groups of English speakers after they had read either the original text or its back-translation. They were instructed to base their estimates upon the essay they had just read. Subjects were systematically assigned to either condition. A total of 86 subjects read the original text and completed the Galileo instrument; 91 read the back-translation and completed the pair comparisons. Data collection took place in the classroom.

Step 6:
Perform Galileo Analysis in the Translated Language

The Galileo instrument was then translated into Arabic and checked through standard back-translation procedures. It was administered to the 50 Arabic speakers after they read the translated text. They were also instructed to base their estimates upon the essay they had just read. Data were collected while students were socializing in the cafeteria.

Results

Step 7:
Analysis of the Differences Among the Versions

In all, 23 (.2 percent) estimates were removed from 9664. No values were removed from the translated data set. The mean reported value for the

original group was 15.81; it was 16.72 for the back-translation and only 6.99 for the Arabic group. Clearly, the Arabic subjects perceived the differences among the scaled concepts to be smaller than did the two English-language groups. This may be due either to the text, which did not differentiate the terms to the same degree, or to the use of smaller numbers by the Arabic speakers. Evidence in support of the latter interpretation may be found by examining the mean reported value for the criterion standard. The Arabic subjects reported that "needs" and "dependency" were only 6.40 units apart, significantly less than the standard of 10.0 ($t = 5.11$, $p < .001$). The original text group reported a value of 10.07, and the back-translation group, 10.53. There was no difference between these two groups ($t = .03$, $p > .05$). To control for the smaller metric, all values were multiplied by 1.55 to adjust them to the same metric as the one used by the original group. All further analyses were based upon these adjusted values.

One difference between the original group and the group that read the back-translation was that there was more variance in the estimates of differences among the pairs for the back-translation. The mean standard deviation for the original was 18.49; this figure was 23.46 for the back translation. The ratio of the variance produced an F of 1.49. If the degrees of freedom are equal to the number of observations (the number of pair comparisons times the sample size), they are 4715, 4923. F is significant, indicating greater overall variance in the back-translation.

This result indicates that there is greater uncertainty about the relations among the key concepts in the back-translation. That is, the translation and back-translation resulted in the loss of information. This result would be expected from information theory (Shannon & Weaver, 1949), which suggests that there is never perfect fidelity in communication. Each time a message is communicated it is distorted to a certain degree. In this case, the code system of the message was intentionally altered twice, once into Arabic and once back into English, causing the loss of information. Since the goal of the translation process is the production of an equivalent document, one without the loss of information, it may prove useful to examine the respective variances of each group to ensure that they are equivalent. This analysis could not be performed with the Arabic group because the use of a smaller standard of comparison is reflected in the variance.

The means of each group were next converted into spatial coordinates. Due to a clerical error only the first ten concepts were scaled. "Feelings" was excluded from the measurement process. The first two dimensions of each data set are graphically represented in Figure 25.1.[2] These two dimensions accounted for 59 percent of the variance for the original group, 54 percent for the back-translation, and 52 percent for the Arabic group. The correlations of the concepts' locations on these axes with the locations of the origi-
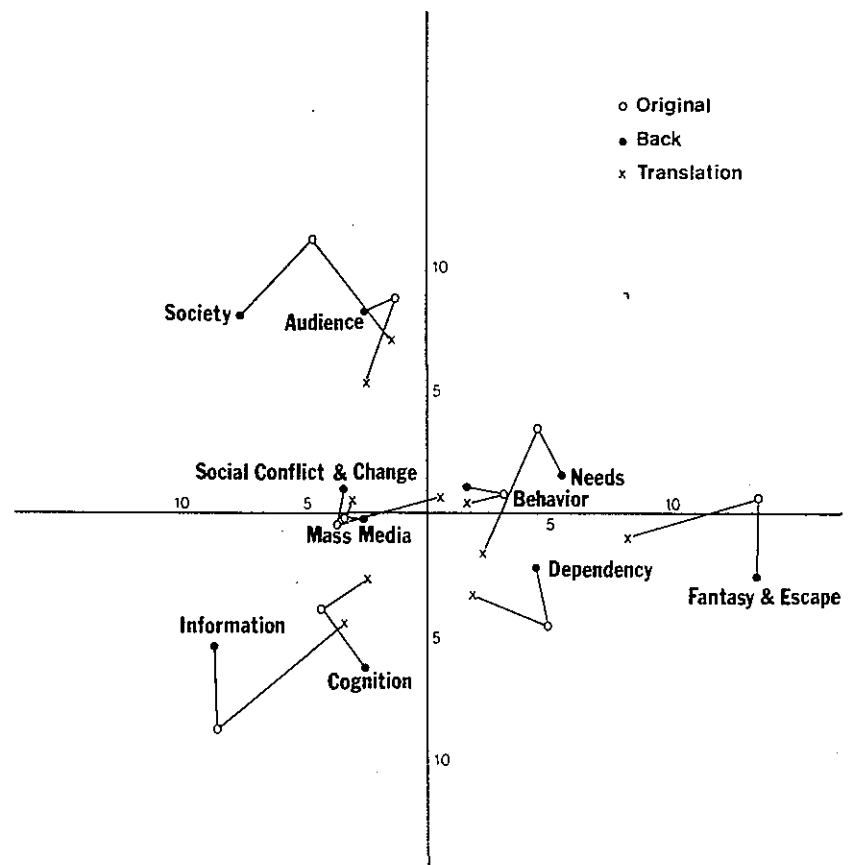
Figure 25.1   Graphic representation for original, translation, and back-translation
groups.

nal group were .98 and .93 for dimensions 1 and 2 for the back-translation
group, and .92 and .99, respectively, for the Arabic group.

All three groups used the same dimensions to evaluate the concepts in
the essay. Dimension 1 differentiates reality from fantasy. The concepts "in-
formation" and "fantasy and escape" are bipolar. Dimension 2 differenti-
ates "society" and "audience" from "information" and "cognition," sug-
gesting a social to mental or psychological dimension. These dimensions
were labeled without statistical verification. Since they are only orthonor-
mal reference vectors, meaning should not be attributed to them unless re-
gression analysis confirms the attribution (Barnett & Woelfel, 1979).

Figure 25.1 reveals that the Arabic group's concepts (the x's) are closer

to the origin than the other two groups. This indicates that they did not differentiate the concepts to the same degree as the other groups. Since these dimensions account for the greatest proportion of the variance, this interpretation is justified.

Figure 25.1 also reveals the sources of the discrepancies among the translations. The length of the lines connecting the equivalent concepts indicates the degree of discrepancy, such that the longer the line, the greater the problem in that concept's translation. These lines are relatively long for "society," "information," "social conflict and change," and "fantasy and escape." Special attention should be given to these concepts when changing the document to improve the translation. The discrepancies are considerably less for "mass media," "behavior," and "audience," suggesting that these concepts were translated with little error. However, since these two dimensions account for only about 55 percent of the total variance, the graphic representation should be viewed somewhat skeptically. Greater emphasis should be placed upon the mathematical description that is presented later.

While there is a high degree of similarity among the spaces, there are systematic differences among the three groups. The traces or sum of the eigenroots of the spatial coordinates matrices indicates that the groups differentiated the key concepts to a greater or lesser degree (Danowski, Stoyanoff, & Barnett, 1977; Stoyanoff & Fink, 1981). The original group's trace was 1052.34; the back-translation, 1097.65; and the Arabic group only, 554.62. This was after the values in the distance matrix were adjusted to account for their use of a smaller criterion standard. The individuals evaluating the text in Arabic failed to differentiate the concepts to the same degree as the readers of the English versions, suggesting that the translated text did not differentiate the key concepts to the same degree as the English versions.

The spatial coordinates of the two translated groups were next rotated to a least-squares best fit, which minimized the departure from congruence with the coordinates of the group that evaluated the original English text (Woelfel, Holmes, & Kincaid, 1979). This made it possible to examine the overall differences between the texts and which key concepts, when translated, were responsible for these discrepancies. The differences between the spaces for each concept and the correlations of the concepts on all the dimensions with the original space are presented in Table 25.1.

Unlike the plot, which describes only about one-half of the variance (that which is attributable to dimensions 1 and 2), the values in Table 25.1 account for the differences on all the dimensions upon which there is variation.[3] As a result, when evaluating the differences between the texts, one should not place too much emphasis on the plot; rather, one should examine the differences across all dimensions (Table 25.1). The correlations allow

Table 25.1
Differences and Correlations Between the Original
Text and the Arabic and Back-Translations

| | Arabic | | Back-Translation | |
	Difference	Correlation	Difference	Correlation
Society	5.25	.95	5.11	.91
Audience	4.94	.90	3.21	.95
Fantasy and escape	4.95	.90	3.21	.98
Social conflict and change	5.32	.69	4.97	.89
Needs	.66	.94	2.43	.96
Cognition	3.65	.97	3.52	.94
Mass media	3.54	.84	1.54	.97
Information	5.15	.89	3.58	.93
Behavior	1.66	.91	2.62	.98
Dependency	3.28	.96	2.34	.98
Mean	3.71	.895	3.25	.949

one to determine if the discrepancies are due simply to the use of larger or smaller values when differentiating the concepts or to a discrepancy in the relations among the concepts. A high correlation and a high difference between groups suggest that the difference is due to scale size, while discrepancies in relations among concepts are indicated by a low correlation and a high difference.

Table 25.1 reveals that the overall discrepancies among the spaces are small, only 32.5 percent to 37.1 percent of the distance between "needs" and "dependency." Overall, the correlations among the spaces are quite high (r = .895 and .949, respectively), indicating that most of the differences between the texts were due to the differentiation of the key concepts. This high level of agreement was expected because the "same" text served as the stimulus for all three groups.

An examination of the individual differences and the correlations reveals that six concepts warrant particular attention in the Arabic translation. They are "society," "audience," "fantasy and escape," "social conflict and change," "mass media," and "information." The correlations suggest that the discrepancies involving "society," "audience," and "fantasy and escape" are largely due to the subjects' inability to differentiate these terms to the same degree as the readers of the English texts. The mean of the pair comparisons involving "society" was 16.53 for the original group and 11.37 for the Arabic group. This difference, 5.20, represents nearly all the discrepancy involving that concept. For "audience," the mean values for the two groups were 15.08 and

12.14, and for "fantasy and escape," 17.31 and 14.57. The differences, 2.94 and 2.74, represent smaller proportions of the total discrepancies that are reflected in their lower correlations, indicating that much of the discrepancy is due to relational differences with other concepts.

"Social conflict and change" was the most discrepant concept between the two conditions. There was 5.32 units difference and its correlation was only .69. Clearly, this concept's meaning was altered during the translation process. In particular, its relations with "audience" and "fantasy and escape" changed 5.19 and 9.17 units, respectively. This accounts for the other portion of the change in these concepts' locations.

According to the translators, the Arabic speakers do not differentiate between "information" and "mass media" to the same degree as Americans. The data bear this out. The mean difference between these two concepts was 8.07 in English and only 6.91 in Arabic. This difference altered the concepts' pattern of relations with the other concepts and produced the somewhat lower correlations across the conditions.

The comparison between the evaluation of the original text and the back-translation revealed slightly greater correspondence than between the original and Arabic treatments. The mean difference was 3.25 compared to 3.71, and the mean correlation was .949 compared to .895. The reason for the greater correspondence is that cultural differences do not affect the results. The process of translation is the only source of differences between the groups.

The primary areas of difference between these two treatments are the concepts "society" and "social conflict and change." The former was 5.11 units different and it correlated .91. The later was 4.97 units different and it correlated .89. The greatest difference between "society" and the two groups was its relation with "information." The difference was 5.46 units and accounted for the discrepancy in the position held by "information" (3.58, r = .93). "Social conflict and change" differences were spread fairly evenly among the other nine concepts, although "fantasy and escape" was 5.63 units different. This suggests that the translation of "social conflict and change" warrants special attention in subsequent translations.

Step 8:
Intervention by the Research to Reduce Discrepancies

The analysis of the differences reveals several points of discrepancy. These areas should be given special attention when the document is re-translated to reduce the discrepancy between the texts. The first problem in the translation is that the Arabic version does not differentiate the concepts to the same degree as the English versions. One way to adjust the Arabic text

is through the use of adjectives and adverbs. Cliff (1959) and Woelfel and Fink (1980) have shown that adjectives and adverbs act as multipliers when attached to concepts. For example, the sentence "The one we focus upon first is the high level of dependence of audience on mass media" could be translated as "The one one we focus upon first is the very high level of dependence of audience on mass media." This should expand the differences within the Arabic space to produce the same differences among the concepts.

In this data set, the concepts of concern are "society" and "social conflict and change." Particular attention should be paid to these concepts when translating the document a second time. What are the relations of these terms to the others in the document? How are they different in the back-translation from the original text? The relations between "society" and "information" and between "social conflict and change" and "fantasy and escape" should be clarified in the translation by using modifiers, or additional words or phrases, to specify the relations as precisely as possible. The discrepancy between the latter pair of concepts involved multiple-word concepts. These have the greatest potential for ambiguity, suggesting that special care may be necessary when translating phrases.

The Galileo analysis may not give sufficient information regarding the domain of concepts evoked by a particular concept. For example, "social conflict and change" may require greater specification than allowed from a single measurement. In that case, it may be necessary to perform additional research by first eliciting those key concepts within the domain of a concept that is particularly difficult to translate. This may be done using interviews by asking subjects in the original language to describe the concept; for example, "Please describe 'social conflict and change' " or "What does 'social conflict and change' mean to you?" As with the pretest, a content analysis of these interviews should reveal those concepts used to define "social conflict and change." At this point, a second Galileo could be conducted to describe the term's relations with the other concepts used to define it. This increase in resolution or power should help guide the translator when making adjustments in the translated text.

Step 9:
Retranslate and Retest

At this point the text should be retranslated with the guidance of the Galileo analysis. The researcher may wish to retest to ensure that the distortion is reduced. The process may be repeated until the discrepancy between all versions of the text is zero. At this point, the translation process has converged and the texts have reached equivalence.

DISCUSSION

This chapter has demonstrated how the Galileo system may be applied to evaluate the accuracy of translation of technical documents. The analysis showed that the translation did not differentiate among the concepts and produced greater uncertainty in the relations among the key concepts. Further, it indicated the sources of difference or error in the translation. In this example, those discrepancies concerned the concepts "society" and "social conflict and change." Procedures to modify the translated text to ensure equivalence with the original text have been suggested, including the use of adjectives and adverbs and simply devoting special attention to the most discrepant terms. This example involved only a single iteration of the translation-evaluation process. The next iteration should use the knowledge of the differences among the original, the translation, and the back-translation to guide adjustments in the translation. Future research should be conducted to see if the use of these methods results in convergence, or equivalent texts.

There are methodological problems with this research that should be pointed out. They concern internal and external validity. Regarding the former, it was assumed throughout this chapter that the only source of error or discrepancy between the texts was the translation process. Other sources of error could have been differences in the environments in which the data were collected and measurement error. The mean percentage of relative error was 13.7 percent for the original group, 15.4 percent for the back-translation group, and 9.8 percent for the Arabic group. This error may have been responsible for the perception of the discrepancies among the conditions. This error is in part a function of the number of observers of the text (the sample size). It could be reduced as a function of the square root of the increase in sample size (Woelfel, Cody, Gillham, & Holmes, 1980).

Another possible source of error was the variance in the identification of the key concepts. "Society" and "social conflict and change," the two most discrepant concepts, were selected by only 19 and 10 observers, respectively, from the group of 31. Problems concerning "society" were raised even prior to the pretest. During the back-translation process, the translators had difficulty agreeing upon a precise set of symbols for that concept in Arabic. This suggests a degree of validity for these procedures.

The methodological problem of external validity arises in the question of the generalizability of this study. What degree of confidence is there in these procedures as a method to determine the accuracy of all translations? This study was done only as an example. There was only a single sample of text being translated to a single language (Arabic), using a small select sample of observers of that language. The majority of this group were Palestin-

ian, a cultural group involved in social conflict and change. It is not surprising that their perception of this term was different than American students. It should be pointed out, however, that this difference was also found between the back-translation and the original, suggesting that the discrepancy may be the result of the translation process. Still, it is unclear how generalizable these results are to all Arabic speakers regardless of the translated document. Future research should be conducted with other texts using other languages to determine the applicability of these procedures to evaluate the quality of translation.

One theoretical issue concerns how much the differences resulting from these measurements reflect cultural differences and how much they involve linguistic differences. These procedures have been widely used both as a measure of culture and as a measure of semantics. In this study, the subjects varied both in language and culture. It was necessary to use the back-translation to control for cultural differences between the Arabic speakers and the American students. Thus any researcher using these procedures should be careful to control for cultural variation when evaluating translation.

Finally, there is the economic issue of marginal utility. These procedures are costly and time-consuming. To use them for every translation may inhibit intercultural communication rather than enhance cross-cultural understanding by creating a backlog and a reduction in the timeliness of the translated message. These procedures should be employed only when the precision of the translation is critical, such as for technical, scientific, or policy documents. Even in these situations, the problem arises as to how many iterations should be performed to ensure accuracy before the cost of analysis outweighs the benefits.

The problem suggests that future Galileo research should not concentrate on the translation of single documents, but should eventually lead to a kind of indexing of relative concepts as reference frames for limited translation situations. These references would include concepts that are found in particular kinds of technical documents. Unlike the static state of dictionary glossaries, these are easily updated and can be applied over time (Barnett, 1979, 1980).

Furthermore, the results of this study would also indicate that accurate premeasurements of the target population could sensitize and direct translators more quickly to the point of equivalence and convergence. Rather than relying on the highly variant skills and intuitive judgments of bilinguals, accurate measurements will give a more precise picture of the two worlds and indicate the best path toward a resolution of the differences. This could effectively cut the time and cost of the process.

SUMMARY

This chapter tested a method to measure translation accuracy using the Galileo method of metric multidimensional scaling, from a convergence theory perspective. It was argued that the equivalence of technical documents should not be determined solely by the judgments of bilingual "experts," but rather by comparison of the relevant semantic/conceptual domains of the text's users in both the original and target languages. The spatial manifolds generated by multidimensional scaling can be compared to determine problems in translation. These can be used to guide an iterative cybernetic process of testing and adjustment of the translation until convergence of congruence between the original document and its translation reaches an acceptable level of equivalence. A test of the proposed methods using English- and Arabic-speaking students was described that showed that the translated texts did not differentiate among the key concepts to the same degree as the original. The translation process produced greater uncertainty in the relations among the key concepts and the analysis identified the primary sources of error in the translation.

NOTES

1. The following section represents a brief summary of Palmer and Barnett (1984). An examination of that source is recommended for a complete discussion of the theoretical issues raised in this chapter.

2. This plot was produced from the coordinate values that resulted after the coordinates of the two translated groups were rotated to a least-squares best fit, which minimized the departure from congruence with the criterion space, the original English text.

3. Generally, the dimensionality of a space is equal to the number of scaled concepts minus one, although in some cases it may be less. In this case, all three texts were evaluated along nine dimensions (Barnett & Woelfel, 1979).

REFERENCES

Ball-Rokeach, S. J., & De Fleur, M. (1982). A dependency model of mass media effects. In G. Gumpert & R. Cathcart (Eds.), *Inter/media: Interpersonal communication in a media world*. New York: Oxford University Press.

Barnett, G. A. (1974, April). *Social system homophily as a function of communication*. Paper presented at the meeting of the International Communication Association, New Orleans.

Barnett, G. A. (1976). *Bilingual information processing: The effects of communication on semantic structure*. Unpublished doctoral dissertation, Michigan State University.

Barnett, G. A. (1977a). Linguistic relativity: The role of the bilingual. In B. Ruben (Ed.), *Communication yearbook 1* (pp. 397-414). New Brunswick, NJ: Transaction.

Barnett, G. A. (1977b). Bilingual semantic organization: A multidimensional analysis. *Journal of Cross-Cultural Psychology, 8,* 315-330.

Barnett, G. A. (1979, May). *On the nature of synonyms: And this little piggie*. Paper presented at the meeting of the Eastern Communication Association, Philadelphia.

Barnett, G. A. (1980, May). *Frequency of occurrence as an estimate for inertial mass: Pigs in*

space. Paper presented at the meetings of the International Communication Association, Acapulco.

Barnett, G. A., & Kincaid, D. L. (1983). Cultural convergence: A mathematical theory. In W. B. Gudykunst (Ed.), *Intercultural communication theory: Current perspectives* (pp. 171-194). Beverly Hills, CA: Sage.

Barnett, G. A., & McPhail, T. L. (1980). An examination of the relationship of United States television and Canadian identity. *International Journal of Intercultural Relations, 4,* 219-232.

Barnett, G. A., Wigand, R. T., Harrison, R. B., Woelfel, J., & Cohen, A. A. (1981). Communication and cultural development. *Human Organization, 40,* 330-337.

Barnett, G. A. & Woelfel, J. (1979). On the dimensionality of psychological processes. *Quality and Quantity, 13,* 215-232.

Brislin, R. W. (1972). Translation issues: Multilanguage versions and writing translatable English. *American Psychologist, 7,* 299-300.

Brislin, R. W. (1974). Back translation for cross-cultural research. *Journal of Cross-Cultural Psychology, 24,* 21-35.

Brislin, R. W., & Sinaiko, W. H. (1973). Evaluating language translations: Experiments on three assessment methods. *Journal of Philosophy and Rhetoric, 57,* 328-334.

Cliff, N. (1959). Adverbs as multipliers. *Psychological Review, 66,* 27-44.

Danowski, J. A., Stoyanoff, N. J., & Barnett, G. A. (1977, June). *A simple-minded view of TV: A look at users' cognitive complexity.* Paper presented at the meetings of the International Communication Association, Berlin.

Ervin, S. M., & Bowers, R. F. (1952) Translation problems in international surveys. *Public Opinion Quarterly, 16,* 595-604.

Ervin, S. M., Osgood, C. E. (1954). Second language learning and bilingualism. *Journal of Abnormal Psychology, 49,* 139-146.

Kincaid, D. L. (1975). *The convergence model of communication.* Honolulu: East-West Communication Institute.

Kincaid, D. L., Yum, J. O., Woelfel, J., & Barnett, G. A. (1983). The cultural convergence of Korean immigrants in Hawaii: An empirical test of a mathematical theory. *Quality and Quantity, 18,* 59-78.

MacNamara, J. (1970). Bilingualism and thought. In J. E. Alatis (Ed.), *Georgetown roundtable of languages and linguistics* (pp. 25-40). Washington, DC: Georgetown University Press.

Nida, E. (1966). A review of *Syntactic Translation* by Wayne Tosh. *Language, 42,* 851-854.

Nida, E. (1975). *Language structure and translation.* Stanford, CA: Standford University Press.

Osgood, C. E., May, W. H., & Miron, M. S. (1975). *Cross-cultural universals of affective meaning.* Urbana: University of Illinois Press.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning.* Urbana: University of Illinois Press.

Palmer, M. T. (1981) *Linguistics and translation.* Unpublished master's thesis, Rensselaer Polytechnic Institute.

Palmer, M. T., & Barnett, G. A. (1982, May). *A method to measure accuracy of language translation.* Paper presented at the meeting of the International Communication Association, Boston.

Palmer, M. T., & Barnett, G. A. (1984). Using a spatial model to verify language translation. In W. B. Gudykunst & Y. Y. Kim (Eds.), *Methods for intercultural communication research* (pp. 129-146). Beverly Hills, CA: Sage.

Rogers, E. M., & Kincaid, D. L. (1981). *Communication networks: Toward a new paradigm.* New York: Free Press.

Shannon, C., & Weaver, W. (1949). *The mathematical theory of communication.* Urbana: University of Illinois Press.

Stoyanoff, N. J., & Fink, E. L. (1981, May). *A metric multidimensional scaling of cognitive structure.* Paper presented at the meeting of the International Communication Association, Minneapolis.

Thurstone (1927). A law of comparative judgement. *Psychological Review, 34,* 273-286.

Wigand, R. T., & Barnett, G. A. (1976). Multidimensional scaling of cultural processes: The case of Mexico, South Africa and the United States. In F. Casmir (Ed.), *International-intercultural communication annual—III* (pp. 139-161). Falls Church, VA: Speech Communication Association.

Woelfel, J., Cody, M. J., Gillham, J. R., & Holmes, R. (1980). Basic premises of multidimensional attitude theory. *Human Communication Research, 6,* 153-167.

Woelfel, J., & Fink, E. L. (1980). *The measurement of communication process: Galileo theory and method.* New York: Academic.

Woelfel, J., Holmes, R., & Kincaid, D. L. (1979, May). *Rotation to congruence for general Riemann surfaces under theoretical constraints.* Paper presented at the meeting of the International Communication Association, Philadelphia.