Is Literacy Acquisition Easier in Some Languages Than in Others?

John Downing

It is frequently asserted that the Chinese logographic writing system is more diffi-
cult to learn than the English alphabetic system. This view seems to be based
chiefly on the belief that the large number of Chinese characters is a heavy burden
on the student. But this may be a misconception for two reasons: (1) there are far
more items to be learned in the English system than is generally recognized; (2) the
sheer number of characters to be learned is not in itself an important psychological
factor. What is more important is the extent of redundancy in the system. Numer-
ous alternatives may conceal the nature of the written code from the beginner and
cause confusion. This and other variables in the writing system may prevent the
child from perceiving that writing and print are indeed "visible language."

Whether it is easier to learn to read and write in one language than
it is in another is just one of several problems which hopefully may be
clarified in the new field of research: "Comparative Reading." This
new research approach is based on the proposition that by making
comparisons between the reading behaviors of people in different
cultures and in varying languages a better understanding will be
achieved of the fundamental psycholinguistic processes of reading
and writing and the manner in which these are learned. It attempts
to draw on the living laboratory of languages and cultures by samp-
ling countries with contrasting linguistic and educational back-
grounds.

As an initial effort to open up this area of research, the present
author conducted the survey which produced, among many others,
the findings reported in this article.

Method

Fourteen nations were slected to represent varying cultural, edu-
cational, and linguistic phenomena. These were: Argentina, Den-
mark, Finland, France, Germany, Great Britain, Hong Kong,

Israel, India, Japan, Norway, the Soviet Union, Sweden, and the United States. For each nation experts were selected to write a report on reading in that country. A copy of an earlier speculative paper (Downing, 1969) on the comparative reading method was given to each of the national authors. This insured that everyone would work toward a common set of aims. In addition, each national author was sent an outline of the probable form of the final report to show how his part would fit into the total project, plus some guidelines for preparing his national report. These indicated that the future readers of *Comparative Reading* would wish to compare one national report with another, and that therefore it would be helpful if each author would consider this when writing. However, the guidelines concluded, "The above list is only intended as a guide to some of the variables which have been mentioned by people interested in comparing reading in different countries. There may be other more important problems which must be understood in the study of the reading of a particular country."

This final instruction stressed the open-endedness of the inquiry, because an important aim was to permit spontaneous responses that might show the differences in cultural priorities from one country to another.

The difficult task of analyzing all the resulting data has been guided chiefly by one of Bereday's (1966) principles for comparative education, that ". . . never-ceasing watchfulness by the observer to control his own cultural and personal biases." An important additional precaution that has been taken is to publish in full (Downing, 1973) all the original data from the fourteen countries so that the reader can interpret them himself.

Results

This comparative reading study has indeed proved fruitful; new light is thrown on a number of different aspects of literacy acquisition. But this article focusses on the important question: Is literacy acquisition in the learner's mother tongue easier in some languages than it is in others? For example, it is frequently asserted that it is harder for Chinese-speaking children to learn to read than it is for English-speaking children. Another popular view in the English-speaking countries is that Spanish teachers of reading have an even easier task in developing literacy in their students.

Conventionally, writing systems are classified into three types. For example, Lado (1957) has three categories "according to the units of language that they represent." These are:

1. *Alphabetic* writing systems in which the characters represent phonemes of the language.

2. *Syllabic* writing systems in which the characters stand for syllables rather than phonemes.

3. *Logographic* writing systems in which the characters represent morphemes—that is, units of form and meaning—rather than units of sound such as syllables or phonemes.

All three of these types were represented in the present study. For example, Great Britain, Israel, and the USSR all have *alphabetic* writing systems, although the actual characters used to represent the phonemes of English, Hebrew, and Russian, respectively, are different. Hong Kong provides a case of a country using a *logographic* system for its Chinese language. Japan generally combines *syllabic* and *logographic* systems.

The general belief is that the order of difficulty in learning these alternatives is from alphabetic (the easiest) to logographic (the hardest) with some uncertainty as to the level of difficulty of syllabic writing systems. For example, Goody (1968) states, "Many scholars of China have seen a connection between a low literacy rate and the use of a non-alphabetic script." He quotes one Chinese writer: "The 'Chinese script' of China is certainly too difficult for the masses and only the gentry class can have time enough to learn it, so that politically and culturally it is an enormous impediment."

What is believed to be the cause of the difficulty in the Chinese logographic writing system? Goody states that "Any system of writing which makes the sign stand directly for the object must be extremely complex." This complexity arises, Goody claims, because Chinese "can extend its vocabulary by . . . making the sign stand either for a more general class of objects or for other referents connected with the original picture by an association of meanings which may be related to one another either in a continuous or in a discontinuous manner. Either process of semantic extension is to some extent arbitrary or esoteric; as a result, the interpretation of these signs is neither easy nor explicit." Therefore, Goody argues, each new character has to be learned "as a separate sign for a separate word."

This means in effect that "a minimum of 3,000 such characters have to be learned before one can be reasonably literate . . . and with a repertoire of some 50,000 characters to be mastered, it normally takes about twenty years to reach full literate proficiency. China, therefore, stands as an extreme example of how, when a virtually non-phonetic system of writing becomes sufficiently developed to express a large number of meanings explicitly, only a small and specially trained professional group in the total society can master it, and partake of the literate culture."

But, apart from primitive pictographs, all writing systems use arbitrary signs. Therefore, Goodys' criticism of the logographic system is concerned essentially with the *very large number of characters* that the literacy learner must acquire. Halle (1968) makes this same point by analogy: "Since the strokes [in Chinese characters] are arbitrary symbols, the writer's or reader's task is equivalent to that of a person trying to remember telephone numbers. And, since in order to read a newspaper one needs to be able to read several thousand words, the person who wishes to read a Chinese newspaper must have memorized several thousand arbitrary stroke sequences. This task is roughly equivalent to memorizing several thousand telephone numbers. . . ."

But in the present comparative reading study the native authors of the national reports on Hong Kong and Japan, which both use logographic systems, do not seem to view them as being exceptionally complex.

Leong, writing on literacy acquisition in Chinese in Hong Kong, actually quotes the same paragraph from Halle in order to stress his disagreement with it: "Although each character has to be learned, the often mentioned reliance on rote memory is overrated. . . . The comparison to memorizing telephone numbers is not quite apt, as chunking and other mnemonic devices can make the task more meaningful and increase one's storage capacity." Leong presents a closely reasoned argument that there is much greater similarity between learning to read in English and Chinese than is commonly supposed. In particular, the Chinese logographs can be analyzed into smaller units that operate systematically, and thus the burden of memorization may be much less than Goody and Halle suggest.

Similarly, Sakamoto and Makita in their report for this com-

parative reading project point out that in Japanese Kanji logographs can be classified according to various radicals. Therefore, students can learn that Kanji are not unsystematic symbols but are constructed from combinations of different radicals.

Thus the effect of having to memorize thousands of characters in Chinese and the somewhat smaller number in the Kanji system of Japanese may have been overestimated. On the other hand, the burden of memorization in the alphabetic writing system has certainly been underestimated. Many parents and teachers in Britain and the United States believe that their children have to learn only twenty-six characters in the roman alphabet. But, in truth, English orthography is much more complex. There are many additional characters that are overlooked by adults who have long since forgotten what it was like to be a beginner in literacy learning:

1. Printed characters take alternative forms: upper-case letters as well as lower-case ones, and some variations on the latter. There are also the varieties in the cursive form of handwriting. Thus, the word *beg* may have the following alternative graphic forms:

beg, BEG, Beg, beg, Beg, beg (etc.)

2. An alphabetic writing system's units are not merely the individual characters. The alphabetic system is primarily a code for phonemes. English has approximately forty phonemic units (varying with dialect), which are supposed to be signaled by the graphemic units of the writing system. Obviously, twenty-six individual letters cannot be enough for the forty phonemes. Furthermore, the variations listed as the first source here do not have any special phonemic function. The English phonemes that do not typically have a single-letter grapheme in the writing system are represented instead by digraphs: *ch, th, sh, ng, oo,* and so on. There would need to be at least forty graphemes to code the forty phonemes if the code were perfectly systematic.

3. In addition there is a great deal of redundancy in the English writing system. Although there are only about forty phonemes, some of them have many alternative graphemes: *ie, y, uy, igh, eye, i ui,* for the single phoneme common to the words *pie, my, guy, high, eye, rind, guide.*

Ellis (1845), the nineteenth-century linguist, analyzed all the alternative ways of printing and writing the forty or so phonemes of English and found more than 2,000 alternative graphemes.

Other alphabetic writing systems vary in the extent to which their total number of graphemes extends beyond the small number of characters in the basic alphabet. English probably is especially prolific in alternative graphemes, but the writing system of Hindi has more than one source of variation (which increase the total number of characters well beyond the fifty-two usually listed as the basic Hindi characters), for example, the alternative vowel symbols to be employed according to position in the word, and the special symbols for consonant clusters.

Therefore the difference between logographic and alphabetic writing systems may have been exaggerated in respect to their difference in number of characters. This may have caused an overestimate of the difference in the length of time needed to learn them. For example, Goody's claim that "it normally takes about twenty years to reach full literate proficiency" in China, assumes that it takes less time in Britain or the United States. But what does "full literary proficiency" mean? One could claim that very few people achieve full literate proficiency in the English writing system, if that phrase includes the ability to spell from memory any word according to the conventions of English orthography. Probably, both in China and in England it is more appropriate to recognize that literacy learning is a lifetime process. The existence of an apparently simple and limited alphabet in the latter country may be misleading to educators in judging the time needed for acquiring literacy.

Nevertheless some notable contrasts in the number of characters to be learned by literacy learners in different countries may exist. For example, Finnish children have very few to acquire, even when alternatives such as capital and lower-case letters are counted. Japanese children, in comparison, have to cope with several alternative writing systems for the one language: Hiragana, Katakana, and Kanji. But, despite the large number of alterative characters to be memorized, Japanese children achieve remarkably well in literacy.

These considerations lead one to question whether the sheer number of characters to be memorized is *in itself* such a significant cause of difficulty in learning to read and write as has been commonly supposed. Extending one's repertoire of logographic characters in

Chinese may not be so very different from extending one's word, phrase, and sentence recognition vocabulary in written English. Both take time and both probably never reach the ultimate ideal level of Goody's "full literate proficiency" in any individual.

Of greater psychological significance in the beginner in literacy learning is the extent to which the features of the array of written characters he meets in his early literacy experiences systematically parallel his past experience of features of spoken language; that is, how far the characteristics of the writing system match appropriately the features of language to which he has been and still is continually exposed. If the writing system is a code of one sort or another for units of speech, the child will understand its symbolic and coding function and mode of operation to the extent that the samples of written language he meets actually demonstrate them. The number of different characters presented to the beginner could have one important influence in this.

If the writing system to be associated with the spoken language is an alphabetic one, then the samples of written language provided for the child ought to reflect its alphabetic nature. Only in this way can the learner understand the characteristics of the decoding and encoding processes of reading and writing. One must hasten to emphasize additionally that if the samples also fail to demonstrate realistically the communicative and expressive functions of decoding and encoding, another, vital aspect of the written code will be missing. Decoding or encoding are meaningless terms if nothing *relevant* gets coded. But, provided that the material is clearly functional for the learner, it ought to reflect as obviously as possible the way in which the code operates. However, if the writing system has many alternative printed and written characters for the same linguistic units, and the initial learning samples contain a variety of these, it is likely to be more difficult for the beginner to perceive the *system* in the writing system. For instance, the English sentence, "I like my pie" quite clearly fits the need to provide beginners with linguistically functional samples, but the basic alphabetic nature of the English writing system is concealed when it is printed in the conventional English orthography:

> *I like my pie.*

All four words contain the same phoneme, which coincides with the whole of the first word "I," but it is represented by four different

graphemes: (1) *I*; (2) *i.e.*; (3) *y*; (4) *ie*. Therefore, this sentence provides no experience whatsoever of the essential grapheme-phoneme basis of an alphabetic writing system.

It is this effect of an abundance of alternative characters with no *easily perceived* logic which makes English orthography difficult for learning to read. Of course, one can restrict the vocabulary presented to beginners by using such sentences as *Dan can fan Nan*, but this sacrifices the vital principle that early decoding must demonstrate the communication function of reading.

In learning to write, too, the existence of alternative written symbols either for the same phoneme or for the same morpheme is likely to be an important source of cognitive confusion in this aspect of understanding the logic of the written code for spoken language. For example, in writing the conventional orthography of English, the child must hesitate when writing a word like "cat." He hears /k/ but does not know whether to write *kat* or *cat* because both *c* and *k* are available. Therefore, he hesitates at this choice point. Actually the selection has no simple phonemic basis. But he does not know this, and therefore, it seems probable that the choice is puzzling to him. He hears no difference between the initial phoneme of words he has met in his reading such as *cap* and *kid*, yet they begin with different letters. He may wonder if he has a hearing defect. Hence the noticeable hesitations at such choice points in his written composition.

This is probably a very important source of puzzlement and confusion. It may cause a significant increase in the amount of initial cognitive confusion in the mind of the beginner who is striving to understand how the structure of the written form of language is related to his past experience and developing consciousness of the structure of speech. Writing systems vary greatly in this characteristic, and a cross-language comparison of hesitations in writing or spelling would provide better evidence on this question.

One empirical investigation provides some evidence that redundant symbols do cause difficulty. Malmquist (1964) compared one group of children in Sweden who learned cursive script in addition to manuscript print in the latter part of first grade (as is conventional) with an experimental group who learned *only* the manuscript print. They did not transfer to cursive script until the end of the third grade. The experimental group's written work was superior in clarity and

legibility during the time they continued to use manuscript print. Also they were better in silent reading comprehension. Furthermore, at the end of third grade after only a few weeks of cursive writing it became superior in quality and equal in speed to that of the control group, which had been using cursive writing for over two years. Thus the reduction in the variety of alternative symbols during the first two and a half years led to the experimental group's superiority in reading as well as writing.

Linguistically, of course, there is no logical mismatch between the units of spoken language and the symbols of written language when the former have two or more alternative symbolizations in the latter. but *experientially* for the literacy learner there may appear to be mismatch in this situation. Malmquist's experimental data strongly suggest that cognitive clarity was readily developed in his experimental group because the superfluous variety of symbols was reduced. The amount of unnecessary "noise" in the stimulus situation was cut down sufficiently for these students to perceive more rapidly the important structural elements of the code and the way they operate. In contrast, the control group pupils were hindered in their groping for cognitive clarity by the extra superfluous variations in the cursive characters thrust on them before they had mastered the manuscript symbols.

Other Variables in the Writing System
The number of characters in a writing system is not the only variable to be considered in exploring the problem of whether one language is easier than another for the native literacy learner. The comparative reading project investigated five other possible variables:

1. The type of lingustic unit coded, e.g. morphemic versus phonological units.

2. The complexity of the individual characters in a writing system.

3. The spatial direction of representing the temporal order of the units of spoken language.

4. The names given to the characters of a writing system.

5. The complexity of the decoding and encoding operation in a writing system as it affects the child's *perception* of regularity.

Each of these variables would require another article in *Visible Language*. Briefly, like numerous other investigations, this comparative reading study has raised many more difficult questions than the apparently simple ones with which it began. However, a sort of answer emerges to the question which is the topic of this article: generalizations about the comparative difficulty of literacy acquisition in Chinese or its simplicity in Spanish, for example, are of very dubious validity. There may be differences in the relative difficulty of literacy acquisition in different languages, but it is a much more complex problem than is usually supposed.

But of greater importance for the future of teaching children to read and write is the final conclusion to our survey of all these variables: "One thing, at any rate, is quite clear from these investigations: the writing system is an important variable in the learning-to-read process. A perceived mismatch between the writing system and the language in which the child is expected to respond is an important cause of difficulty in the early stages of learning to read and write."

Writing and print become visible language *only when the child perceives them as visible language*. Writing systems contain many potential sources of confusion which may cloud the child's recognition of their function and operation. It is an important task for all who contribute to this field to find better ways in every language of presenting writing and print to children, so that they may more readily arrive at a clear perception of them as visible language.

REFERENCES

Downing, J. Comparative reading: a method of research and study in reading. In Figurel, J. A. (ed.). *Reading and Realism*, Newark, Del.: International Reading Association, 1969.

Downing, J. *Comparative Reading*. New York: Macmillan, 1973.

Ellis, A. J. *A Plea for Phonotypy and Phonography*. Bath: Isaac Pitman Phonographic Institution, 1845.

Goody, J. *Literacy in Traditional Societies*. London: Cambridge University Press, 1968.

Halle, M. Some thoughts on spelling. In Goodman, K. S., and Fleming, J. T. (eds.). *Psycholinguistics and the Teaching of Reading*. Newark, Del.: International Reading Association, 1968.

Lado, R. *Linguistics Across Cultures*. Ann Arbor: University of Michigan Press, 1957.

Malmquist, E. *Overgang from textning till vanlig skrivstil*. (With summary in English). Stockholm: Kungl, Skoloverstyrelsen, 1964.