EYE MOVEMENTS: From Physiology to Cognition
J.K. O'Regan and A. Lévy-Schoen (editors)

SYNCHRONIC DESCRIPTION OF PRESENT-DAY WRITING SYSTEMS: SOME IMPLICATIONS FOR READING RESEARCH +

Daniel Holender

Laboratoire de Psychologie Expérimentale
Université Libre de Bruxelles
Brussels, Belgium

First, the characteristics of present-day writing systems are illustrated by describing how Chinese, Japanese, Korean, Hebrew, Arabic, and Hindi are conveyed by script. This is followed by a discussion about the nature of the linguistic units that are, or that should be represented by alphabetic scripts. The last part of this chapter is a brief description of the different lines of research bearing on how the reading processes are modulated by the properties of the writing systems.

INTRODUCTION

The first of the two sections of this chapter is aimed at describing the main present-day writing systems as a function of the relationship between the written symbols and the linguistic units they represent. From the standpoint of a cognitive psychologist interested in reading processes, it is almost natural to emphasize a synchronic rather than a diachronic approach to writing systems. This is because it is unlikely that in reading a person brings to bear much knowledge about how the script has evolved to reach its present state of perfectness or imperfectness. The second section elaborates on the first in dealing with orthographic representations and in outlining how the cognitive psychological (and neuropsychological) approach to reading became interested in investigating the reading processes in different writing systems.

Anybody interested in writing systems would be struck by the lack of recent general surveys of this topic. Standard references are Diringer (1968) and Gelb (1963) in English and Cohen (1958) and Février (1959) in French. One should not be misled by these publication years that, except for Cohen, correspond to revised editions of books basically summarizing the state of the art before World War II or a little later. All these books reflect outstanding scholarship in the historical and descriptive aspects of writing systems, and Cohen (1958) and Gelb (1963) should even be credited for going far beyond this exclusive point of view. Cohen's (1958, Chapter 10) treatment of orthographies in many alphabetic languages has plenty of insightful distinctions and observations still worth reading for those more concerned with linguistic aspects of writing. Gelb (1963) was even more ambitious in aiming at laying the foundations of a science of writing. In the course of doing so, Gelb made several personal theoretical choices, some of which were judged highly controversial by many of his peers. This point should be kept in mind because, during the last 30 years, Gelb's book has become the standard reference in the field. As such, it exerted a nonnegligible influence on the cognitive psychology of reading, for instance in shaping some of the ideas developed by Gleitman and

+Preparation of this chapter was supported by the Research Council of this University and by the Ministry of Scientific Policy (Action de Recherche Concertée "Processus Cognitifs dans la Lecture").

Rozin (1977 ; Rozin & Gleitman, 1977) in two very influential papers. I shall come back to this point in the second section. There is no recent, general survey with the scope and comprehensiveness of those I have just mentioned (or of Menninger, 1969, in the related topic of number writing) and, sadly, it is likely that our new scientific habits will no longer encourage such books to be written.

With respect to theory and terminology, two chapters in a book edited by Haas should be mentioned ; one by Pulgram (1976) is concerned with typologies of writing systems, and the other by Haas (1976) deals with the basic options in designing writing systems. Very recently, a new book by Sampson (1985) was published that deserves to become a standard reference in the field, at least among cognitive scientists. From the vantage point of the linguist, Sampson succeeds in integrating typological, historical, and psychological aspects of writing systems to a degree never reached before.

Space limitation prevents an in-depth discussion of terminology and definition. To a large extent, I shall adopt Sampson's hierarchical (1985, p. 32) classificatory scheme, which is a satisfactory synthesis of concepts developed by Gelb (1963), Haas (1976), and Pulgram (1976).

The first major distinction is between semasiographic and glottographic visual systems of communication. A semasiographic message conveys meaning directly without mediation of the spoken language. Semasiographic messages were generally vague when they were first used in primitive societies, but modern semasiography as it is used in specialized domains such as road signs or mathematical notation conveys highly specific, precise messages. The cardinal feature of a semasiographic message is that it does not translate into a single sentence of the spoken language. In contrast, a glottographic message represents an actual or a potential spoken utterance. In his initial broadest possible definition of writing as "a system of human intercommunication by means of conventional visible marks" Gelb (1963, p. 12) did not distinguish between semasiography and glottography. But he went on to restrict writing to glottographic systems, coming close to the conception of Bloomfield (1933, p. 21) who claimed that "writing is not a language, but merely a way of recording language by means of visible marks". A minor addition to these definitions is needed to avoid any visual message, such as nonverbal meaningful gestures or finger spelling, to be considered writing ; namely, the visible marks must be made permanent on some traditional support (stone, wood, paper, leather, etc)[1].

Glottographic systems being the only ones to provide full writing, any further distinction in the classification will be restricted to them. Recording language by means of visible marks requires establishing some kind of isomorphism between the phonic medium and the graphic medium. This could theoretically be achieved in many different ways, but actualizations have taken only four different forms. Mapping between the two media has been implemented by segmenting the spoken utterance into words, morphemes, syllables, or phonemes and by associating these units with conventional graphic symbols henceforth referred to as graphemes[2]. It is convenient to consider an intermediate level in the classification in which glottographic writing systems based on meaningful units (words or morphemes) are subsumed under the logographic heading and those based on meaningless, phonological units, either syllables (syllabic systems) or phonemes (alphabetic systems), under the phonographic one. Three points should be made to further specify some terminological distinctions.

1. The term logography has replaced the older term ideography that became inappropriate as soon as one recognized the glottographic essence of writing systems. What is represented in writing is not the concept itself (or the idea), but the word standing for the concept. To keep strictly in line with etymology, the term logographic should designate only writing systems whose units are words. We could then use the term morphographic for those systems representing morphemes. This would force us to find a new header for replacing the logographic one, but little would be gained in doing so. Besides, it would not help us much in specifying exactly

what is represented in the Chinese script, the only one to be extensively based on a logographic principle. As we shall see, Chinese graphemes stand for monomorphemic words and the notion of plurimorphemic word is not as distinct from that of syntactic construction as it is in English, for example.

2. Number writing is semasiographic, not glottographic, because symbols used in the positional notation do not translate a spoken utterance (e.g., 3,125 stands for three thousand, one hundred and twenty-five). There is nevertheless a widespread tendency among cognitive psychologists, including myself in another article (Holender & Peereman, 1985), to consider single digits as logograms. This is not necessarily an appropriate terminological choice as suggested by Edgerton (1941) in a brief, insightful article. The point raised by Edgerton is that even single digits do not always have an exact correspondance in a spoken utterance (e.g., 2 is not pronounced in reading *2nd* ; word order is reversed in reading *£3*). Hence, from a formal point of view, it is better to classify digits under the semasiographic heading in all circumstances.

3. An alphabetic script establishes a relation between graphemes and abstract units of the spoken language called phonemes, not between graphemes and speech sounds. If the latter were the case, an alphabetic script would be equivalent to the International Phonetic Alphabet (IPA), which is never the case. Phonetic differences such as that between the aspirated [t] in [top] and the unaspirated [t] in [stop] are not represented by the orthography unless they have a functional contrastive value in the language ; that is, unless they differentiate between phonemes. This is not to say that phonetic features are never represented at all, but they are neither systematically represented nor used as basic units in any phonographic writing system. A similar claim can be made as regards suprasegmental phonology (e.g., stress, intonation) that is poorly, if at all, represented in writing.

The following conventions will be used. A word between double quotes refers to a linguistic form irrespective of the medium. An italic word refers to the written form. As usual, phonetic representations are between brackets and phonemic ones between slant lines, but I shall keep on using the Latin alphabet, supplementing it with verbal descriptions if necessary, instead of using the IPA symbols.

PRESENT-DAY WRITING SYSTEMS

Nonalphabetic Scripts

Chinese. The standard references are Kratochvíl (1968) for the language and Alleton (1976) for the writing system. As regards writing, shorter treatments by French (1976), Leong (1973), and Wang (1973) provide us with excellent introductions to the topic.

Chinese is not a unitary language ; it can be divided into several groups of dialects all belonging to the Sino-Tibetan language family. Mandarin is the most widespread group of Chinese dialects of which one variety spoken in the area of Peking is now considered Modern Standard Chinese in Mainland China. What should be understood by this concept is probably well captured in the following quotation from Kratochvíl (1968, p. 21) : "the language used today by educated speakers of Peking dialect which most speakers of other Chinese dialects consider as the 'correct' form of oral communication and in whose favour they adjust their own speech behavior". Yué dialects are spoken in Hong Kong and in the area of Canton and also by the majority of the Chinese people living in the United States because their ancestors generally came from the Canton area (Wang, 1973). In Taiwan, people speak dialects of the Mín group.

Typologically, Chinese is an isolating language. Many words are free, invariable root morphemes that are never inflected. There is also a limited number of bound

morphemes never used in isolation, but serving as affixes associated with root morphemes. Plurimorphemic words are often dimorphemic. In this latter case words consist of a rigid arrangement of either two root morphemes or of a root morpheme and a bound morpheme. Kratochvíl (1968, Footnote p. 73) writes that "MSC [Modern Standard Chinese] morphology is up to a point petrified earlier syntax". The boundary between syntax and morphology is therefore not very easy to draw in Chinese.

Chinese syllabic structure is relatively simple, which entails that only a few hundreds of syllables exist in the language. Though ancient Chinese probably had a repertory of over a thousand of syllables, modern Mandarin has only a stock of about 420 such elements (Alleton, 1976 ; Leong, 1973). Allowing V to stand for a single vowel, a diphthong or a triphthong, only V, CV, VC, and CVC syllables are represented in modern Chinese. This basic syllabic stock is nearly four times as big (1,380 syllables according to Diringer, 1968, p. 64) because almost every syllable can be pronounced with one of four distinctive tones[3]. In Chinese, a tone change (e.g., /mā/ with a high vs. /mà/ with a falling tone) is as distinctive as a consonant change (/mā/ vs. /pā/) or a vowel change (/mā/ vs. /mī/).

The cardinal structural feature of Chinese is that almost every possible syllable is a morpheme and almost every such morpheme is also a word. Even in allowing for widespread homophony--about 80 % of the monomorphemic words have at least one homophone (French, 1976, p. 105)--the basic stock of nearly 1,400 syllables is not large enough to yield all the words required by the language. This entails that many words are plurimorphemic, often dimorphemic (hence disyllabic) as was already stressed above.

Chinese graphemes, generally called characters, map onto speech at the morphemic level. Hence, strictly speaking, the Chinese script is morphographic. However, as point 1 in the *Introduction* entails, using the logographic label is more convenient in our classificatory scheme. Aside from a few exceptions (mainly bound morphemes), there is a one-to-one correspondance between characters and monomorphemic words. Compound words are therefore represented with the appropriate number of characters, each one corresponding to one constituent morpheme. Let us summarize what has been described thus far by stating that in Chinese there is almost a one-to-one relationship between the elements of any two of the four following sets : syllables, morphemes, monomorphemic words, and characters.

Structurally, characters can be classified into two main groups, one consisting of simple characters and the other of compound characters. The make-up of simple characters is a variable number of strokes whose configuration cannot be analysed into functional subcomponents. Many of them have a pictographic origin that is now lost through simplification and stylization. A few of them are diagrams concretizing more or less abstract concepts (e.g., one, two, or three horizontal strokes standing for the numerals 1, 2, and 3). There are about a thousand simple characters in use at the present time.

Compound characters are of two types. There is a small subset of characters composed of two or three simple characters and whose meaning is derived from the meaning of the subcomponents. However, most of the compound characters (probably about 80 % of the characters in use), are phonetic compounds. One such character consists of two parts ; one, called the radical, suggests the meaning of the character, and the other, called the *phonetic*, suggests its pronunciation. There are 214 radicals (now reduced to 189 in Mainland China), most of which are simple characters that can be used either in isolation or, in some cases, as *phonetics* in other characters. *Phonetics* are often simple characters, but some of them are themselves phonetic compounds. One can identify 1260 such *phonetics* (Alleton, 1976, p. 36).

Two qualifications should be made about phonetic compounds. First, the radical may sometimes provide an accurate indication of the meaning of a character, but often the semantic relation between the radical and the character is very loose and even, in some cases, misleading. Second, a similar situation prevails with respect to the

relation between pronunciation of the phonetic part of such a character and pronunciation of the character itself. Sometimes homophony or near homophony (the tone is generally neglected in building these associations) is the rule, sometimes great differences between pronunciations exist, mainly because words once pronounced the same have followed different paths of phonetic transformation.

Whether simple or compound, characters always occupy roughly the same square area in a text. There is no space between characters, hence no visual cues demarcating plurimorphemic words, but Western punctuation marks are now commonly used to separate clauses and sentences. Chinese is written either in columns to be read from top to bottom and from right to left or in rows to be read from left to right.

Chinese morphographic script is well suited for writing the invariable monosyllabic morphemes of the Chinese language. One of its main advantages is to provide an easy means of disambiguating homophones in writing. This could hardly be realized by using an alphabetic script (even allowing for a lot of irregularity in the grapheme-phoneme correspondance such as in the French homophones *mer*, *mère*, *maire*, for example) because homophony is so widespread in Chinese that it allows complete short homophonous sentences to be written (a trick sometimes used for greeting purposes, see Sung, 1979). One of the main drawbacks of Chinese script is the difficulty of writing foreign proper names (Sampson, 1985, pp. 166-167). Such words are first given a Chinese syllabic rendition that is then transliterated using characters having the appropriate pronunciation, their meaning being generally disregarded (e.g., *Marx* = *mǎkèsī*). There is no particular subset of Chinese characters used as a syllabary, but the choice of characters playing the role of *phonetic loans* is smaller than the full repertory. In analysing the writing of 200 foreign names appearing in Hong Kong newspapers (seven different newspapers in a four-day period), Godwin (1979) found 267 different characters used for representing 193 different syllables, of which 52 were associated with more than one character.

There is no doubt that mastering Chinese reading and writing requires a considerable amount of rote leaning. There is, however, a widespread tendency in the West to exaggerate the burden imposed by the Chinese script. Besides, the saliency, in terms of linguistic awareness (see the second section, below), of the monosyllabic morphemes represented by Chinese graphemes probably makes the writing system easy, if not quick, to learn.

How many characters should a Chinese person know to become a reasonably proficient reader ? Different sources provide an estimate of 2,000 to 3,000 characters corresponding to the standard of elementary education. In reviewing thirty years of teaching studies taking place before 1949, Ai (1950) reported that an average student knew about 2,000 characters by the end of his four-year primary courses. Similarly, in Hong Kong, children learn from 500 to 600 characters each year, thereby acquiring a total of about 3,000 to 3,500 characters after six years of primary school (Leong, 1973). According to Leong, only a quarter of the schooling time (plus a fair amount of daily homework) is devoted to this topic. In Mainland China, there is an official list of 2,421 characters everyone should know. Current dictionaries often comprise less than ten thousand characters[4] (Ai, 1950 ; Alleton, 1976). Different sources agree in estimating that an adult of average education reading newspapers probably knows about 5,000 characters, which implies an ability to read about three times as many words (Ai, 1950 ; Alleton, 1976 ; French, 1976) because many words are plurimorphemic.

Japanese and Korean. The standard reference about Japanese language is Miller (1967). A special issue of *Visible Language* edited by Seeley (1984a) is devoted to some aspects of the Japanese writing system. In the psychological literature, Morton and Sasanuma (1984) provide a good sketch of Japanese writing and Taylor (1980) does the same for Korean ; Sakamoto (1980) gives an excellent description of *hiragana*.

For historico-political reasons, a large stock of Chinese morphemes were incorporated into both the Korean and the Japanese spoken languages. These languages belong neither to the same family (Korean is an Altaic language, but whether Japanese also belongs to this family is still controversial) nor to the same typology as Chinese (Korean is agglutinative and Japanese inflexional). Words, and even root morphemes, are often plurisyllabic. These languages differ, however, in the complexity of their syllabic structure. Japanese has only a little over 100 open syllables of the types CV and V and some CV syllables containing a glide. In contrast, Korean has several thousands of syllables of the following types : V, VC, CV, CVC, and CVCC. Writing both of these languages using exclusively Chinese characters standing for their morphemic value would be inadequate because affixes, grammatical inflexions, and function words could simply not be represented. The solution adopted in both Japanese and Korean consists in writing the roots of content words (nouns, adjectives, and verbs) with the corresponding Chinese characters and to supplement this incomplete system with either a syllabary (Japanese) or an alphabet (Korean) to represent all the rest. Such is the principle, but, as we shall see shortly, its implementation is much more complex, especially in Japanese.

To begin with, let us describe the phonographic part of these systems. Japanese's two *kana* scripts, *hiragana* and *katakana*, are complete syllabaries, either of which could be used to write any Japanese word, whether of native or foreign (including Chinese) origin. There is a one-to-one correspondance, similar to that between upper and lowercase letters in our alphabet, between the graphemes in the two *kana*.

There are 46 basic *kana* graphemes ; five corresponding to the vowels /a, i, u, e, o/, one to the nasal phoneme /n/, and the rest to open CV syllables of which the consonants are /k, s, t, n, h, m, y, r, w/ and the vowels any of the five vowels just mentioned. The syllables /yi, ye, wi, wu, we/ are not represented ; some because they never existed in the language, others because they are no longer in use today. Twenty-five additional syllables are represented by adding one of two diacritics to the upper right side of some of the basic characters. A small circle transforms /h/ into /p/. Two ticks transform /k, s, t, h/ into their voiced counterparts /g, z, d, b/. Thirty-three syllables containing a glide (e.g., /nyo/ or /kyu/) are written with two characters, one corresponding to the syllable beginning with the appropriate consonant followed by an /i/, and the second being a small version of the character for /ya/, /yo/ or /yu/ (Hence, *nyo* is *ni* + small *yo* and *kyu* is *ki* + small *yu*). There is one other small character used to indicate gemination of any consonant.

The Japanese grapheme-syllable correspondance is regular, except for at least five cases (see Sakamoto, 1980). One such case is that in the middle of a word, the phonetic value of [n] changes into [m] if the following phoneme is /m/, /b/, or /p/, but the same grapheme is used in both cases. Two other cases are that the *kana* for the vowels /u/ and /i/ should be pronounced as if they were standing for /o/, and /e/ when they are used to prolong the sound of syllables ending by /o/ and /e/. Two further cases involve the grammatical particles /wa/ and /e/ that are written with the graphemes usually standing for /ha/ and /he/ ; this being an example of historical orthography.

Although it is quite convenient from a typological point of view to consider the two *kana* as syllabic scripts, this terminology is only roughly appropriate. The point is that *kana* symbols stand for pronunciation units of approximately the same duration. These phonological segments are called moras. In Japanese, moras are not always coextensive with syllables, the most prominent exception being the sound [n], which is a mora, but not a syllable. This distinction is best captured by an example. A word such as "ningen" (human being) has two syllables (/nin/ and /gen/), but four moras (/ni/, /n/, /ge/, and /n/) ; it is accordingly written with four *kana*.

The Korean phonographic component could not be syllabic because syllables are too numerous. Therefore, an alphabetic principle is used. Only 24 basic letters are needed to represent the phonemes of Korean (Taylor, 1980) ; 14 standing for the consonants and 10 standing for the vowels[5]. These alphabetical symbols are always grouped into syllabic blocks containing as many letters as there are phonemes in the syllable (from 1 to 4) except that the absence of initial consonant in a V or a VC syllable is indicated by a small circle. Letters are organized in both the horizontal and vertical dimensions of a syllabic block according to simple rules. Knowing these rules and the basic alphabetic symbols allows a person to read any syllabic block and hence any plurisyllabic word written phonographically (not including a logographically represented root).

The logographic component of the Japanese script consists of *kanji* characters. A *kanji* character is a borrowed Chinese character standing for a Japanese native root morpheme having roughly the same meaning as that originally represented in Chinese. If this were all, the system would be relatively simple. Complications arise from two sources. First, many Chinese words were incorporated into Japanese, forming the Sino-Japanese vocabulary. This implies that the Japanese spoken language which, as any language, is of course able to express any current concept with native words, often has Sino-Japanese synonyms or near-synonyms to express the same concepts. Sino-Japanese words are naturally represented with the original Chinese characters. It follows that most characters have at least two different "readings", one called the *kun*-reading, corresponding to the Japanese native morpheme, and the other, called the *on*-reading, corresponding to the Sino-Japanese form of the same morpheme. Second, there are three different layers of Chinese words in Japanese, each layer originating in a different period of borrowing (roughly, 6th, 8th, and 14th century A.D. : see Miller, 1967, Chapter 3). Chinese pronunciation rules were different during each of these periods. Even though the vocabulary borrowed during each of these three periods is more or less disjoint, there are, however, some characters having two or even three different *on*-readings, the correct reading being determined contextually (see Sampson, 1985, Chapter 8 for more details).

The pronunciation of a Sino-Japanese word consists in the Japanese rendition of the corresponding Chinese pronunciation at the time of borrowing. Relative to Chinese, tone is always suppressed, which implies that homophony is extremely high in monosyllabic Sino-Japanese words (less than 300 different forms exist), and sound modifications are frequent because Japanese phonology is different from that of Chinese. All this implies that the phonetic part of a compound character, which can of course play no phonetic role at all in the *kun*-reading of the character, is probably not playing a much greater phonetic role in the *on*-reading(s) of the same character either.

Pronunciation of the *kanji* part of the script is contextually determined. Generally, isolated *kanji* are given a *kun*-reading, whereas *kanji* occurring in compound words are given an *on*-reading. Which particular *on*-reading is appropriate when there is more than one possibility can be determined by subtle meaning distinctions conveyed by the compound. *kanji* completed by *kana* indicating inflexions are always given a *kun*-reading. No such complications exist for the logographic part of the Korean script because there is only one possible pronunciation which is quite close the Chinese one (Taylor, 1980).

The Japanese script is, therefore, a mixed script combining logographic and phonographic representations. Content root morphemes are written in *kanji* ; inflexions, function words, affixes, and other grammatical particles are written in *hiragana* ; foreign, non Chinese, loan words are written in *katakana*. Also, Arabic numerals are now commonly used to write numbers, and uppercase Latin letters appear in Japanese texts for common international abbreviations.

Like Chinese, Japanese is written either in columns to be read from top to bottom and from right to left, or in rows to be read from left to right. Word boundaries

are not indicated but the more complex visual configuration of *kanji* compared to *kana* graphemes provides a good visual cue for root morpheme identification. Also, a transition from *hiragana* to *kanji* must represent a word boundary because there are almost no prefixes in Japanese (Sampson, 1985, p. 186). The Western full stop and comma are now commonly used (see Twine, 1984).

The main advantage of the logographic part of the Japanese script is to allow for disambiguation of homophones, which is even more necessary in Japanese than in Chinese. Another advantage pointed out by Backhouse (1984, p. 223) is that the scientific vocabulary is transparent because it is built on Chinese roots which, once represented by the corresponding characters, could evoke their more or less synonymous meaning in native Japanese. For these and other reasons, it is extremely unlikely that the Japanese would ever drop the logographic component of their script. Assuming they do nevertheless abandon it, there is no reason (at least, no linguistic reason), in view of the simplicity of the syllabic structure and of the associated *kana* syllabic script, that Japanese should ever switch to an alphabetic script. Notice, however, that Miller (1967, p. 98) states that "it was the nature of the script which was borrowed, rather than the structure of the language to which is was now adapted, which determined the syllabic nature of Japanese orthography".

Before World War II, a Japanese reader should probably have known as many (5,000 or so) *kanji* characters as a Chinese reader. The situation has changed since 1946 because the *kanji* repertory has been limited to an official list of 1,850 characters (see Seeley, 1984b for a history of changes in the Japanese script since 1900), of which 881 (996 since 1971) are supposed to be learned by the end of primary school (Sakamoto & Makita, 1973). A revised list published in 1981 now contains 1,945 characters, of which 737 have only *on*-readings, 40 have only *kun*-readings and 1,168 have both *on* and *kun*-readings (Backhouse, 1984). The official list of *kanji* characters represents the minimum educational standard and newspapers are requested not to use characters outside the list. This is not so in scientific or literary texts so that, even now, Japanese educated readers probably know over 3,000 characters. In North Korea, the use of Chinese characters has been abandoned completely, whereas in South Korea 1,300 such characters are still used in combination with the alphabetic-syllabic script (Taylor, 1980).

Alphabetic Scripts

The alphabetic principle is used in almost all the written languages of the world save for those we have just discussed. Ideally, an alphabet should contain as many graphemes as there are phonemes in the spoken language and a one-to-one correspondance should be established between graphemes and phonemes. Although this ideal is almost never reached completely, it is often approximated quite closely. It is worth distinguishing different types of departures from this principle because they are not all equally damaging to it. One type of departure consists in using polygraphs. A polygraph is a combination of two or more graphemes of the repertory standing for one single phoneme. As long as these graphemic groups are consistently mapped onto the same phonemes, their use only slightly violates the basic principle, if at all. A little more damaging is the existence of consistent, but contextually determined values of some graphemes (e.g., the value of *c* before *e* and *i* vs. its value before *a*, *o*, and *u* in English or in French). Finally, the most damaging departure is achieved when the one-to-one grapheme-phoneme correspondance is abolished for reasons unrelated to contextual letters. In this case, one grapheme (or polygraph) may represent several phonemes or one phoneme may be represented by more than one grapheme, or both. Some reasons why such a violation of the basic alphabetic principle can occur will be examined in the second section of this chapter.

In what follows I shall provide some broad historical information (based almost exclusively on Diringer's 1968 account) to explain how alphabetic writing spread

all over the world. There is general agreement about the fact that the alphabetic writing principle evolved only once in history. There is, however, much controversy about the stages (and the influences involved) of the process, taking several centuries, which led to the first complete alphabet : the North Semitic alphabet (called West Semitic syllabary by Gelb, 1963). One of the oldest famous inscriptions using this alphabet, that of Ahiram's coffin, is now dated as stemming from the early tenth century B.C. (rather than from the late thirteen century B.C. as believed before, see e.g. Gelb, 1963, p. 132 ; Harden, 1971, p. 108). The North Semitic alphabet, with its 22 letters representing only consonants, is the prototype of all the alphabets in the world. It is convenient to distinguish three main branches in the spreading of alphabetic scripts ; these are the Aramaic, the Brahmi, and the Greek branches. Aside from their different geographical distribution, which is irrelevant for our purposes, these three groups of alphabets differ mainly with respect to how fully and by which graphical device they represent vowels.

Aramaic branch : Hebrew and Arabic alphabets. Starting in the seventh century B.C. and lasting for more than a thousand years, Aramaic, a Semitic language from Syria, became the *lingua franca* of the Near East, extending as far as West India. Aramaic was written with a graphically modified version of the North Semitic alphabet. The Aramaic script is at the origin of both the classical "square Hebrew" alphabet and the Arabic alphabet. Numerous classical Hebrew inscriptions can be found in the second century B.C. and the earliest Arabic ones appeared in the fifth or sixth century A.D.

The principle of Semitic writing is to represent fully the consonants and to leave the vowels either unmarked (originally) or incompletely marked as is the case in most Hebrew and Arabic texts today. Two features of the Semitic language structure are probably responsible for the fact that a script based on such a principle can work. First, the phonological make-up of a root morpheme is generally not concentrated into one unitary speech segment. Rather, being embodied in the consonantal skeleton of a word-size utterance, this make-up is discontinuously spread throughout the spoken word. Second, the vowels interspersed within such a consonantal skeleton are used as inflexional patterns indicating, for instance, part of speech, verb tense and mood, and so on. Hence, a given consonantal pattern (often comprising three consonants) corresponds to a whole group of words related in meaning. In the case of an isolated word, providing the vowels (as is done in spoken language) would generally unambiguously indicate which word is intended. Conversely, in reading a text, contextual syntactic and semantic information together with the consonantal skeleton of a given word would, theoretically, unambiguously determine which word it is, thereby indicating which vowels to pronounce in the case of oral reading (see Barr, 1976).

Actually, reading a Semitic text completely devoid of vowels is not as easy as the foregoing description might erroneously suggest. It would require considerable grammatical knowledge and a cumbersome, long semantic and syntactic analysis would be necessary to reach a correct interpretation. For these and other reasons two ways of indicating vowels have been devised. One system, by far the most widespread, consists in providing partial vocalic information by using *matres lectionis* ; that is to say, letters of the repertory that could stand either for consonants or for long vowels according to the context. This system, though generally helpful, may nevertheless constitute a source of ambiguity in its own right because some letters play a dual function. The second system for specifying vowels consists in adding diacritics to the consonants. This optional practice is generally referred to as "pointing" and a text could therefore be "pointed" (containing both *matres lectionis* and diacritics) or "unpointed" (containing only *matres lectionis*). These diacritics are used not only for supplying full vocalic information, but also for completing missing consonantal information. There is no such practice as partial pointing ; hence, an unpointed text lacks both the vocalic and the consonantal information conveyed by diacritics.

Rosén (1977) provides a linguistic analysis of the contemporary Hebrew spoken language. As regards writing, Barr (1976), and especially Rabin (1977) are good sources and Navon and Shimron (1984) give us an excellent psychologically (from an information-processing point of view) oriented description of the main features of present-day Hebrew script.

Unpointed Hebrew is written with the 22 original consonantal letters derived from the Aramaic alphabet. Some letters stand for two different phonemes. Conversely, some earlier phonemic distinctions now lost in the spoken language are still orthographically represented ; hence, Hebrew orthography is historical.

Modern Hebrew has only five vowels (six in some pronunciations) /a, e, i, o, u/. The distinction between long and short vowels is now completely lost in speech, but not in writing. Rabin (1977) distinguishes between two types of unpointed spelling ; these are grammatical spelling and full spelling. Grammatical spelling uses three of the 22 basic letters as *matres lectionis* for supplying partial vocalic information about (former) long vowels exclusively. In this system, the letters *v* and *y* could stand either for their consonantal value or for one of two vocalic values (*v* stands for /o/ or for /u/ and *y* stands mostly for /i/ but sometimes also for /e/). A third *matres lectionis*, *h*, always indicates a vowel at the end of a word (generally /a/, sometimes /e/, and /i/ in a few cases) and a consonant elsewhere. Notice that grammatical spelling is completely standardized and consistent, also it is defective within its own logic in not being able to represent the long /a/ except in final position, and that it offers no graphical means of distinguishing between the vocalic and the consonantal roles of *v* and *y*. Full spelling adds two features to grammatical spelling. First *v* is now used for representing the (former) short vowels /o/ and /u/ and *y* for representing the short /i/. Second, the vocalic role of these two letters is now distinguished from their consonantal role by doubling the letters in this latter case. Notice that the use of full spelling is inconsistent and not standardized and that vowel representation is still ambiguous (many-to-one vowel-letter mapping) and incomplete because long /a/, short /a/, and short /e/ cannot be specified.

Most adult literature is unpointed. Pointing is used mainly in the Bible, children's books, and poetry. Children start learning to read with pointed texts, but they rapidly switch to unpointed ones. The pointing systems consists of a set of diacritics (placed above, below, or inside the consonants) aimed at supplying almost all the information missing in unpointed spelling. Pointing is quite successful in achieving this goal, but new problems may arise from the fact that the system is plethoric. For instance, there are 14 diacritics for indicating five vowels ; hence, here again Hebrew orthography pays its tribute to history. Diacritics are also used for distinguishing between cases in which the same letter could stand for two different consonants, for distinguishing between the vocalic and the consonantal value of *matres lectionis*, and for a few additional roles that will not be described here.

Hebrew is written from right to left, words are spaced, and letters are not very discriminable (square, noncursive shapes with few ascenders and descenders). Five letters have two different graphical versions, one being used only in initial and medial position and the other only in final position.

The basic principle of Arabic writing is the same as that of Hebrew. There are, however, some differences bearing mainly on details. The alphabet contains 29 letters standing for consonants only. Three of these letters are *matres lectionis* that are used consistently in unpointed spelling for specifying the three long vowels of Arabic /a, u, i/. Contrary to Hebrew, there is a one-to-one relationship between *matres lectionis* and long vowels in Arabic writing. In pointed spelling, the three other vowels of Arabic (short /a, u, i/) are marked by diacritics. There are two more diacritics, one used to indicate the absence of a vowel, and the other for consonantal gemination. On the whole, the Arabic pointing system is less complex than that of Hebrew due partly to the fact that all conso-

nants are represented by basic symbols and partly to the small number of vowels.

Contrary to Hebrew, Arabic writing is extremely cursive. The basic shape of some letters are the same, but these are differentiated by using one, two, or three diacritical points. These diacritics are not optional and should not be confounded with those used in the pointing system described above. About three fourths of the letters have four different shapes according to whether they stand alone, or in initial, medial, or final position in a word. Arabic is written from right to left and words are separated by spaces.

The Brahmi branch : Devanagari alphabet. All the modern Indian alphabets derive from the Brahmi alphabet. This alphabet, appearing in the seventh century B.C., was probably adapted from the Aramaic alphabet (this hypothesis is favoured by Diringer, 1968, p. 262, though he admits that only the idea, not the script itself, could have been borrowed). Indian alphabets and their foreign offshoots in South East Asia, Indonesia, and Tibet are numerous. One of them called Devanagari (earliest inscriptions in the seventh century A.D.) plays a very important role in India because it was used for writing both Sanskrit (the historical literary language) and Hindi (the official language of modern India).

I took most of my information about Devanagari from Diringer (1968), Février (1959), and Oommen (1973). Devanagari has 48 letters : 34 consonants (plus a few additional borrowed consonants according to Oommen, 1973), 10 vowels, and 4 diphthongs. A consonant letter never has a pure consonantal value, but indicates the consonant followed by a short /a/, a very frequent vowel in the spoken language. This short /a/ is not pronounced in final position and becomes a *schwa* in unstressed syllables. The short /a/ is therefore indicated by default, whereas a vowel different from it is always fully specified by a small sign appended above, below, or to the side of the consonant. The value of a consonant letter accompanied by a vowel sign is therefore an open CV syllable whose vowel replaces the short /a/ of the basic letter. Two additional features are important. First, an initial vowel could not be represented in the system described thus far. This problem has been solved by designing vowel letters that are used only in initial position. Second, consonantal clusters are represented by grouping two or even three consonantal letters, which sometimes involves putting one letter on top of the other and modifying letter shape and integrity. There seems to be a large number of different consonantal clusters in Devanagari because, according to Février (1959, p. 348), 256 corresponding graphic patterns were available for printing at the "Imprimerie Nationale" in Paris. One diacritic indicates nasalization of the following vowel and another indicates less than full phonetic value of a letter (Oommen, 1973).

Devanagari is written from left to right and words are now separated by spaces. One feature common to all letters is a horizontal top line from which the other letter features hang. In addition to spacing, words are also individualized by the fact that the top line of the component letters is continuous.

One should not be misled by the syllabic appearance of the script. The principle that is applied is genuinely alphabetic, both consonants and vowels being completely represented by individual graphemes and a particular vowel sign having the same shape whatever the consonant to which it is appended (see Barr, 1976, pp. 72-74 for a similar claim about the Ethiopian script whose origin is the South Semitic alphabet rather than the Brahmi alphabet). Another point is that alphabets of Brahmi origin are not more defective in representing vowels than alphabets derived from Greek ; the difference between them being more of graphic than of linguistic import. Compared to those of Western alphabets, vowels (except in initial position) of Indian alphabets are smaller than consonants and not interleaved between consonants on the main line of text.

The Greek branch. The Greeks adapted the North Semitic alphabet in order to write their own language sometime during the first centuries of the first millennium B.C. (earliest inscriptions in the eighth century B.C.). All the alphabets used in European languages (and later in most languages spoken in the U.S.S.R.) are derived from Greek by mediation of either the Latin or the Cyrillic alphabet. The Greek's important innovation was the assignment of five unused Semitic letters to represent the vowels. They also added some letters to represent missing phonemes, ending up with the 24-letter classical Greek alphabet in the fourth century B.C. This alphabet was later modified by the Romans, who used 23 letters. Three letters were added in the Middle Ages, leading to the modern Latin alphabet of 26 letters used for all the languages of Western Europe and also by the Slavonic people of Roman Christian faith. The other Slavonic people, those of Greek Orthodox faith, have adapted the 43-letter Cyrillic alphabet that was designed on the Greek pattern in the ninth century A.D.

It is noteworthy that the Latin alphabet does not contain enough letters to represent all the phonemes of most West European languages. Because of the widespread reluctance to add new letters, the solution has often been to use polygraphs and to modify the values of some letters by adding diacritical marks. In contrast, the Cyrillic alphabet contains enough letters (sometimes, even too many letters) for a one-to-one grapheme-phoneme correspondance to hold true in most Slavonic alphabets derived from it. Also, after 1917, dozens of previously nonwritten languages in the U.S.S.R. were provided with alphabets containing enough letters to represent univocally each phoneme (this was first achieved by means of a modified Latin alphabet and then by switching to the Cyrillic alphabet during the Forties). As a consequence these alphabets often contain many more letters than we are used to. Amongst the 54 such alphabets listed by Diringer (1968, pp. 378-381), 37 have between 30 and 39 letters, 9 between 40 and 49 letters, and 8 between 50 and 73 letters.

Lack of space prevents me from describing specific scripts (see Cohen, 1958, Chapter 10 for many examples) originating from the Greek alphabet in any detail but some specific problems raised by some of them will be discussed in the next section.

ORTHOGRAPHIES AND READING RESEARCH

Orthographic Representations

In most languages the orthographic conventions are more complex than is suggested by simply describing the basic relation between graphemes and the linguistic units they generally stand for. Some reasons why this is so have already been touched on in the preceding section, albeit implicitly.

One complication arises from the mixture of different principles within a single text, the most extreme case being the mixture of phonograms and logograms in Japanese or in Korean. Such a mixture is not completely unknown in Western orthographies because a small number of nonalphabetic symbols such as £, $, %, &, and Arabic numerals are sometimes encountered in alphabetically written texts. As was mentioned in the *Introduction*, numbers, even single-digit numbers, are best considered semasiograms and the logographic nature of some other symbols is also questionable (e.g., $ and £ in $3 and £7).

What really complicates orthographic conventions, however, is the fact that the graphemes do not always keep their primary function. In Chinese, a basically logographic unit may sometimes take a phonographic value as is the case when it plays the role of a *phonetic* in a compound character or that of a *phonetic loan* for writing foreign proper names. Similarly, a string of letters may preserve the visual make-up of morphemes at the expense of loosing the phonemic values of some letters (e.g., *heal* vs. *health*) or words pronounced the same (homophones) may be

differentiated orthographically (e.g., French *ver, verre, vers, vert,* and even so *vers* is a homograph with two meanings). One can, therefore, hardly specify what is represented in an orthography in terms of a single principle.

A good starting point for discussing these matters may be found in two citations, one from Gelb (1963) and the other from Chomsky (1970), offering as contrasting views as possible about what an ideal orthography should represent. Gelb reflects the linguistic *Zeitgeist* of the first half of this century in claiming that :

> what we should look for is a system of writing combining the exactness of the IPA alphabet with the formal simplicity of a shorthand system. (p. 246)

Whereas Chomsky (1970) believes that :

> an optimal orthography ... would be one that has as close a correspondance as possible, letter-to-segment, to the abstract lexical form. Such an orthography leads directly to the semantically and syntactically significant units, abstracting away from all phonetic properties that are determined by general rule. (p. 12)

Today, nobody would maintain that an orthography should be phonetic, but the underlying conception according to which an orthography should provide regular, consistent spelling-to-sound mapping rules is still widespread. As should be clear from the *Introduction*, this formulation is acceptable only if 'sound' is taken to mean any speech sound realizing the same phoneme. Gelb's original position in the above citation can accordingly be reformulated in the following terms without greatly altering its spirit : An optimal orthography should establish a regular, consistent correspondence between the graphemes (or grapheme strings) and the phonemes of a language (this will be referred to as Gelb's amended position from here on).

Chomsky's position stemmed from the generative theory of English phonology he developed with Halle (see Chomsky & Halle, 1968, pp. 49-50). The aim of this theory is to account for variations in the phonetic forms of words (e.g., "sign"-"signal", "divine"-"divinity", "courage"-"courageous") by postulating underlying invariant lexical representations and a set of ordered phonological rules for converting these representations into the appropriate phonetic representations. Chomsky believes (see above citation) that an optimal orthography should establish a point-by-point relation between letters and units of the underlying lexical representation, which is roughly what is achieved by English ; hence Chomsky's claim that "conventional orthography, in the case of English (or any other language of which I have any knowledge), is remarkably close to optimal in this sense" (Chomsky, 1970, p. 7).

It should be further pointed out that Chomsky's position rests also on the, needless to say, controversial assumption that some of the formal distinctions made by linguists have a "psychological reality" in the mind of a competent speaker-hearer/writer-reader. Chomsky postulates the psychological reality of only two types of representations--the lexical and the phonetic representations--departing from many other theorists in denying the existence of a phonemic level of representation. The phonological representation is composed of units more abstract than phonemes because they are defined in terms of both phonetic and morphemic properties (see Gleitman & Rozin, 1977). These units may be called morphophonemes and the lexical representation may be considered morphophonological, even though Chomsky avoids this terminology.

One supplementary hypothesis worth mentioning is that the lexical representation is assumed to be the same for every speaker of a language, even for speakers of

different dialects, and to remain the same for long periods of time in spite of phonetic changes affecting the spoken language. This accounts for the adequacy of a single, stable orthography for representing different dialects and different stages of evolution of the spoken languages.

Recently, Mattingly (1984, see also I.Y. Liberman, A.M. Liberman, Mattingly & Shankweiler, 1980) has taken the argument further in attempting to demonstrate that orthographies are basically morphemic. According to this view, whatever the basic linguistic units represented by the graphemes, the ultimate goal of any orthography is to keep morphemic shape constant. This goal is, of course, automatically achieved in Chinese because it is a morphemic script. In phonographic scripts, the only way to fulfill this requirement would be to use an orthographic representation mapping the underlying morphophonological lexical representation as closely as possible. As explained above, this is basically what Chomsky believes orthographies do, but Mattingly introduces an important additional distinction, that between shallow and deep orthographies. A shallow orthography maps graphemes onto units corresponding to a broad phonetic transcription, units that are almost equivalent to phonemes as defined in Gelb's amended position. A deep orthography maps graphemes onto the morphophonemes of the underlying lexical representation. Languages may also be characterized by the shallowness or deepness of their phonology according to whether a small or a large number of phonological rules are needed to transform the lexical representation into the phonetic one. Of course, the deepness-shallowness dimension of both orthography and phonology is best conceived of as continuous, but, for the sake of simplicity, let us pursue the argumentation as if it were dichotomous. Combining deep and shallow orthographies with deep and shallow phonologies yields four possible outcomes. The crux of Mattingly's argumentation is that genuine shallow orthographies do not exist, or seldom exist. For him, it is most likely that what may look like a shallow orthography is determined by the shallowness of the phonology ; that is, by the fact that the lexical representation actually represented in the orthography is close to the phonetic representation.

Granting that one cannot hope to distinguish between shallow and deep orthographies in languages with shallow phonology, refuting Mattingly's hypothesis would require showing that languages with deep phonology do not necessarily have deep orthography. Here follow four lines of evidence casting doubt on the plausibility of the hypotheses of Chomsky (1970) and Mattingly (1984).

1. Historically, the basic goal in designing most alphabetic orthographies, including that of English, was to represent phonemes ; that is, to fulfill Gelb's amended position. The same tendency is pervasive when new orthographies are now adapted to unwritten languages (Berry, 1958, 1977 ; Tauli, 1977). Assuming that it is unlikely that in all these cases orthography is shallow because phonology is shallow, why then design phonemic orthographies if morphophonemic ones are more appropriate ?

2. Very few orthographies are explicitly designed to represent deep morphophonological structures, though O'Neil (1972) documents the recent (19th century) introduction of a morphemic, alphabetic orthography for Faroese. When orthographies happen to be deep, reforms tend to make them shallower. For example, the original Semitic scripts provided close to ideal morphophonemic orthographies, but they started loosing some of their perfectness in this respect when the decision was taken to indicate some vowels by means of *matres lectionis*. Moreover, full writing of the predictable vowels is even supplied (by pointing) when correct interpretation must be guaranteed (e.g., in the Bible or in poetry). Similarly, in Japanese, the often predictable change of a morpheme's initial unvoiced consonant into its voiced counterpart when the morpheme is the second of a compound (e.g., /kana/ becoming /gana/ in /hiragana/) was not marked in the orthography before, but it is now always indicated by a diacritic. In this respect Mattingly's (1984, p. 16 ; I.Y. Liberman et al., 1980, p. 147) assertion that this predictable change is not marked in *kana* is incorrect, probably because it is based on a

confusing description of Martin (1972, p. 93). It is clear that the *kana* component of the Japanese script is shallow, not deep as Mattingly believes (see Miller, 1967, pp. 194-196 ; Sampson, 1985, p. 186). This is, of course, not to deny that the Japanese orthography is highly morphemic, but it is so thanks to its logographic, not to its phonographic, component. Notice, however, that the voicing of the initial phoneme of morphemes occupying the second position in compounds is no longer always predictable in modern Japanese (indeed, /kana/ gives /hiragana/, but also /katakana/). Hence, one could argue that this unpredictability is responsible for the shallowing of an otherwise deep orthographic convention, just as the predictable vowels are now indicated in Turkish orthography perhaps because many borrowed words do not respect the rules of Vowel Harmony (Mattingly, 1984, p. 16). The present argument would, of course, be considerably weakened if this hypothesis were correct. But one could equally argue that if a deep orthographic principle were so desirable, the use of a shallower orthography could well be confined to borrowed words that, in any case, can generally be easily differentiated from native words. Indeed, English orthography, though containing many borrowed words, remains a deep orthography as regards Vowel Shift in the native vocabulary (I.Y. Liberman et al., 1980, p. 147).

3. Spelling-pronunciation, the tendency to pronounce words as they are written, is a phenomenon opposite to that described in point 2. When spelling and pronunciation become discrepant with respect to grapheme-phoneme mapping rules, instead of writing what is actually pronounced by changing the orthography, one can equally pronounce what is actually written. The spelling-pronunciation phenomenon is well documented by Levitt (1968, 1978) for English, French, German, Italian, and Spanish. It seems, therefore, that the underlying intuition of the speaker-hearer/writer-reader is that spelling should be phonemic (Gelb's amended position). Moreover, spelling-pronunciation may affect directly the form of the lexical representation and, more generally, orthography may shape the lexical representation even when spelling-pronunciation does not occur (Skousen, 1982). It is not clear to me how such phenomena could be reconciled with Chomsky's position.

4. Chomsky's claim about English orthography has been criticized on several grounds (Francis, 1970 ; Lotz, 1972 ; Tauli, 1977 ; see Sampson, 1985, Chapter 10 for a reappraisal). Even if the description of English orthography is more or less acceptable, one should be very cautious in interpreting Chomsky's (1970, p. 7) sibylline claim about "any other languages" he knows because the suggested generalization, though not studied systematically, is probably false. To take an example, it is often asserted that Schane (1968), in analysing French phonology and morphology, reached a conclusion similar to that of Chomsky for English as regards French orthography. This is hardly so. It is true that Schane (1968, p. 16) wrote that "French spelling, to a large extent, is highly morphophonemic" but he did so in the context of the analysis of French elision and liaison and he could hardly have concluded the same (indeed, he did not, but he eschewed the problem completely) after analysing the vowel and the verb system of French. Moreover, even in the context of elision and liaison Schane (1968, p. 17) immediately qualified his assertion by providing an example for which "the form is representative of a stage of derivation between the morphophonemic and the phonemic or phonetic levels". It is clear that admitting that intermediate levels of phonological derivation may be consistently represented in the orthography amounts to refuting Chomsky's position. One final point that deserves consideration is that although it is unlikely (but this should, of course, be investigated) that French phonology is shallower than English phonology as regards at least some cases of vowel alternation, it is nevertheless clear that French orthography is much shallower than English in transcribing many of the resulting phonetic (phonemic) distinctions (contrast English *divine-divinity, human-humanity* with French *divin-divinité, humain-humanité,* for instance). Similarly, Spanish, replete as it is with productive alternations such as *sentar-asiento, contar-cuento,* and so on, is certainly not particularly shallow phonologically though quite shallow orthographically.

Of course, impressionistic statements hardly constitute decisive arguments. Yet, in the light of the four lines of evidence just mentioned, I would conclude that one can hardly maintain the general claim (Mattingly, 1984) that orthographies aim at preserving the visual shape of morphemes, regardless of phonemic distinctions, which in alphabetic scripts implies that orthography should map onto the underlying morphophonological representations of morphemes (Chomsky, 1970 ; Chomsky & Halle, 1968). Even English orthography, which seems often to confirm this principle, also relies on other principles for representing language.

In a way, it is surprising that it took so long to recognize that English orthography is not unsystematic, but that it establishes regularities at different levels, which implies that it has to compromise between several incompatible tendencies. Bolinger (1946) and Vachek in the thirties (see Vachek 1945-1949, 1959, 1973 for English accounts) were among the first to point to the fact that English orthography often preserves the visual appearance of morphemes disregarding the phonemic changes if necessary (Bolinger, 1946 used the term "visual morphemes"). Venezky (1970b ; 1970c) has systematically studied the structure of English orthography, concluding that "English spelling-to-sound correspondences can be described in terms of patterns of graphemic, morphemic, syntactic, and phonotactical processes" (Venezky, 1970c, p. 127). Venezky (1970a), recognizing that it is advantageous for an orthography not to rely on a single principle, suggested that in designing new orthographies one should ponder the advantages of not relying exclusively on regular grapheme-to-phoneme correspondence rules.

It seems clear to me that the only reasonable research strategy would be to pursue Venezky's enterprise in English and in many other languages, coming up with a quantitative description of the proportion of cases in which the orthography is based on a phonemic (Gelb's amended position), a morphophonemic (Chomsky's position), and on other intermediate principles. Even though each of these orthographical principles is likely to be represented in at least some circumstances in most languages, alphabetically written languages should nevertheless differ greatly in the frequency of usage of each of these. Correlating the orthographical descriptions with structural descriptions of the languages (at least from phonological, morphological and morphophonological points of view) should help understand why this is so, which in turn could help generate interesting psychological hypotheses about how reading is achieved.

Reading Research

Lack of space prevents even thinking of establishing subtle distinctions or convergences between different reading research areas. To describe how a growing interest for reading in different writing systems evolved, I shall nevertheless adopt a trichotomic partition of the cognitive approach, distinguishing between a psycholinguistic, an information-processing, and a neuropsychological stream of research on the reading process. Two collections of chapters provide us with a wide coverage of this topic ; one was edited by Kavanagh and Venezky (1980) and the other by Henderson (1984). I shall conclude by some prospective considerations about the role eye movement monitoring may play in this research.

Psycholinguistic approach to reading. People working in this framework have been influenced to various degrees by Chomsky's theoretical positions as regards language acquisition and by his hypothesis described above about what should be represented in orthography. Related to these positions is the concept of linguistic awareness, whose theoretical aspects are developed by Mattingly (1972, 1979, 1984) and whose practical implications for research on reading development and reading disabilities can be well appreciated, for instance, in Gleitman and Rozin (1977 ; Rozin & Gleitman, 1977) and in I.Y. Liberman (1982 ; I.Y. Liberman et al., 1980).

The basic ideas are the following. First, analysing and describing how language is conveyed by script is of paramount importance. Second, reading is less natural

than listening to speech. Reading is, therefore, a derived activity that, unlike speech processing, requires linguistic awareness, that is, an ability a person has to reflect upon his own linguistic activity. Third, most scripts represent phonological units, hence phonological awareness is the kind of awareness required for reading most scripts.

Linguistic awareness is necessary for learning to read because as I.Y. Liberman (1982, p. 58) explains : "any reader or writer must, at least, be able to abstract from the utterances of a language exactly those units that the orthographic characters represent". The problem is that not all these units are equally accessible, or equally salient. Though it is quite easy to abstract words or morphemes, it is more difficult to abstract syllables, and even more difficult to abstract phones[6]. Hence, in these terms the speech units idealized by alphabetic scripts are the most difficult to abstract from the phonic substance, probably because there are no such discrete units individualized in the speech stream. Though containing three phones a spoken utterance such as [big] is not composed of three successive units of sound equivalent, say, to three successive piano notes. Acoustic cues corresponding to each phone overlap in time through the whole syllable and abstracting phones from the speech stream requires a complex perceptual process, of which even the end-product is not always readily available to consciousness (A.L. Liberman, Cooper, Shankweiler & Studdert-Kennedy, 1967).

The ability to consciously abstract phones from the speech stream is nevertheless necessary to understand how an alphabetic script works. Conversely, this ability develops as a byproduct of learning to read an alphabetic script as the following results show. Suppose the task is to delete the first phone of spoken utterance such as /rapin/, thus answering with /apin/. It has been shown that children just beginning reading instruction (Alegria & Morais, 1979), adult illiterates (Morais, Cary, Alegria & Bertelson, 1979), and adult Chineses (Read, Zhang, Nie, & Ding, 1984)[7] reading logograms but not *pinyin*, all perform poorly ; whereas children after a few months of reading instruction, adults who became literate during adulthood, and adult Chineses knowing *pinyin*, all perform well (see Content, 1984 ; 1985 for a review of this research area).

One issue raised by Gleitman and Rozin (1977, Rozin & Gleitman, 1977) though somewhat tangential to the study of reading, is nevertheless relevant to the topic of this chapter. These authors surmised that the ontological development of the child may parallel the cultural development of humanity as regards the difficulty in abstracting the phonemes, adding a new example to the already long list of cases in which ontogeny is thought to recapitulate philogeny (see Gould, 1977, Chapter 5, for many such cases). The cultural part of the claim stemmed from an uncritical acceptance of Gelb's (1963) main thesis about the history of writing.

Gelb's thesis is that without borrowing, a writing system would always evolve by passing through three stages : logographic, syllabic, and alphabetic. Although a script may remain at an intermediate stage of development, it cannot skip stages. It is because he firmly believed in this thesis that Gelb denied the existence of consonantal graphemes in Egyptian hieroglyphs and in the Semitic scripts probably deriving from it. Gelb therefore considered these consonantal graphemes to be syllabic, not alphabetic, because an alphabet could not have evolved from a logography without first passing through an intermediate syllabic stage. Edgerton (1952) forcefully criticized this thesis in claiming that the alphabetic nature of some Egyptian consonantal hieroglyphs is uncontroversial among egyptologists and that Gelb's position is a pure example of *petitio principii*. Moreover, it is extremely unsatisfactory, at least in the dominant framework of segmental phonology, to consider a grapheme standing for a consonant followed by any vowel to be syllabic to the same extent as a true syllabic grapheme standing for a consonant followed by a specific vowel. Notice, however, that Gelb's claim that Semitic scripts are syllabic would have found a much better echo in the polysystemic approach to phonology advocated by Firth (1948) ; but this conception never played a prominent role in linguistics (see Lyons, 1962) and did not influence Gelb's

thinking at all. Gelb also believed that not reaching the alphabetic stage is necessarily a mark of backwardness for a writing system. Again, Gelb's lack of linguistic considerations is responsible for his failure to appreciate how suited Chinese and Japanese scripts are to the structure of the languages they represent[8].

Information-processing approach to reading. Information-processing psychologists are interested in describing the nature of the mental representations involved, for instance, in recognizing, matching, or memorizing some stimuli. Various experimental techniques were developed during the late fifties and the sixties that were aimed at tracing the time-course of the flow of information and at probing the nature of transient intermediate representations used for processing. Starting in the early seventies a series of already classical experimental paradigms, and some new ones too, were adapted to fit the requirements of research on visual word recognition. Four main questions are currently investigated ; these are : (1) Is word recognition holistic or mediated by the identification of the constituent letters ? (2) What is the role of sentence context on word identification ? (3) What is the nature of the information used to access the mental lexicon, and related to it, how is pronunciation of a word achieved ? (4) What is the nature of the post-lexical memory code ? (See Henderson, 1982 for a very cogent discussion of the information-processing approach to reading).

Almost all this work has been carried out with native speakers of English, which implies that one cannot distinguish between those processing characteristics that are general and those that result from the idiosyncrasy of English orthography. There is, therefore, a growing body of research on word processing in alphabetic orthographies shallower than English, such as Serbo-Croatian (see Turvey, Feldman & Lukatela, 1984, for a review) or deeper than English, such as Hebrew (e.g., Bentin, Bargai & Katz, 1984). Comparison between pointed and unpointed Hebrew is of course of paramount importance in such an approach (e.g., Navon & Shimron, 1984) as is contrasting the effect of orthography on the same word processing tasks (e.g., English vs. Serbo-Croatian in Katz & Feldman, 1983). Thus far, the work on Chinese and Japanese have been carried out mainly in the framework of the investigation of hemispheric specialization (see Leong, Wong, Wong, & Hiscock, 1985 ; Peereman & Holender, 1985, for recent data and critical appraisals of this approach), but there is also some work on postlexical memory storage (e.g., Tzeng, Hung, & Wang, 1977). A good review of the information-processing approach to different orthographies can be found in Hung and Tzeng (1981).

Cognitive neuropsychological approach to reading. People working in this area use brain damaged patients showing different reading deficits as a source of information to build and corroborate functional information-processing models. Several different types of acquired dyslexias are now well documented in patients reading languages written in the Latin alphabet (mainly English and French), but Japanese patients showing a dissociation between the ability to read *kana* and *kanji* words are also extensively studied (Sasanuma, 1980 ; 1985). The cognitive neuropsychological approach to reading and its relation with the information-processing approach is best appreciated in the already classical *Deep dyslexia* book edited by Coltheart, Patterson, and Marshall (1980) and in its just published offspring, *Surface dyslexia* (Patterson, Marshall, & Coltheart, 1985). Paradis, Hagiwara, and Hildebrandt (1985) provide a similar account for Japanese reading.

Eye movements in reading. However justified is the interest for single word processing, one should nevertheless want to study the reading process in more natural settings than the usual laboratory situations. Eye movement monitoring offers itself as an ideal technique because it is both an on-line indicator of processing and a noninvasive methodology. Since the development of modern eye movement monitoring techniques in the mid-seventies, research has concentrated mainly on determining, for instance, the size of the perceptual span, the amount of information extracted from the parafovea, and how saccades are controlled. Relatively few studies have used eye movements as an indicator of ongoing linguistic processes in reading (with some prominent exceptions, e.g., Just & Cartpenter, 1980), but

the situation is rapidly changing and new experimental paradigms and measurement techniques are now developed by investigators having a strong psycholinguistic orientation (e.g., K. Ehrlich & Rayner; 1983, S.F. Ehrlich & Rayner, 1981; Frazier & Rayner, 1982). A parallel interest in monitoring hand movements in Braille reading also has arisen recently (see Bertelson, Mousty & D'Alimonte, 1985; Mousty et Bertelson, 1985). A collection of chapters edited by Rayner (1983) provides us with a good overview of research on eye movements in reading.

Thus far, very few comparative studies of different orthographies have been undertaken with eye movement techniques. Yet, we already know that Hebrew-English bilinguals have opposite asymmetric functional visual fields (area of text from which useful information is extracted during a fixation) in reading Hebrew and English (Pollatsek, Bolozky, Well, & Rayner, 1981) and that Chinese-English bilinguals (living in California) read Chinese faster in horizontal than in vertical texts (Sun, Morita, & Stark, 1985).

CONCLUSION

Ever since the publication of *Language by ear and by eye* (Kavanagh & Mattingly, 1972) the interest for comparative studies in reading of different orthographies has been growing considerably. Some progress along this line has been made but major breakthroughs await more extended descriptions of how language is conveyed by speech, how it is conveyed by script, and how they are related. More specifically, we need the analysis of languages at the level of phonology, morphology, morphophonology, and syntax that is already provided by linguists, to be more extensively correlated with a fine-grained analysis of how the associated orthographies represent these different levels of description, which we still lack to a large degree.

Having these descriptions would allow us to assess which combinations are actually implemented in a two-dimensional space defined by different levels on the orthographical and phonological deepness-shallowness dimensions and how the reading process is affected by such variables. In the meantime, research has exploited the few cases in which a given language can be extensively represented in two different orthographies. Prominent examples of this approach concern the *kana-kanji* comparison in Japanese (e.g., Paradis et al, 1985), the pointed-unpointed comparison in Hebrew (e.g., Navon & Shimron, 1984) and the possibility of mixing two alphabets in Serbo-Croatian (e.g., Katz & Feldman, 1983). Along this line, one is certainly eager to find a group of Chinese people equally proficient in reading Chinese written in *pinyin* and in the logographic, traditional orthography, a possibility that is perhaps much less remote than in the past.

NOTES

[1]This definition of writing is still defective because it prevents considering the Braille tactile alphabet as a writing system. Lyons (1977, pp. 68-69) solved the problem by distinguishing between the medium in which the language is realized and the channel along which it is communicated. Writing is realized in the graphic medium that is usually communicated by visual marks along the visual channel, but a tactile message (change of channel of communication) isomorphic to the visual message still belongs to the graphic medium. Uncontracted Braille, which establishes a one-to-one correspondance between tactile patterns of dots and the letters of the written language, is therefore no less writing than the usual visible marks.

[2]I use the term grapheme to designate a class of graphs representing the same abstract written unit. For example, B, b, *b*, and **b** are different graphs realizing the same grapheme. English and French are both written with 26 graphemes (disregarding diacritics and punctuation marks). This view is similar to that of

Venezky (1970c, pp. 47-49), but not to that of Sampson (1985; p. 25) who adopts a closer parallel between the graph-grapheme distinction and the phone-phoneme distinction. I find this latter position unsatisfactory as it obliges one to consider upper and lowercase letters as different graphemes because they play different, contrastive orthographic roles.

[3]Four diacritics are used to indicate tone in *pinyin*, the official alphabetic system used to transcribe Chinese. These diacritics, one to be placed above the vowel of each syllable, are ¯, ´, ˇ, and `, corresponding to the high, rising, low, and falling tones, respectively.

[4]The number of entries in a dictionary is always indicated in terms of number of characters. Even though I lack the precise, explicit information, I take this to mean that only monomorphemic words are listed and that compound words never stand as entries. If compound words are indicated at all, they should appear under the heading of one of their constituent morphemes (e.g., *huǒchē*, meaning train, should be found either under *huǒ*, meaning fire, or under *chē*, meaning vehicle).

[5]The Korean alphabet is praised for being the only one in which some letter features portray the shape of the main articulators involved in the pronunciation of the corresponding sound. Korean is, therefore, the most subphonemic alphabet ever designed. This remarkable achievement of fifteenth century scholars is nevertheless immaterial to the main concern of this chapter because it is extremely unlikely that this characteristic of the script plays any functional role in the reading process.

[6]Chinese monomorphemic words should be extremely salient psychologically because of their coextensiveness with syllables that are themselves very clearly demarcated phonologically, the sets of possible initial and final consonants being almost disjoint (Sampson, 1985, p. 146). However, words are not always psychologically salient. For example, many Spanish common nouns of Arabic origin were generally not separated from the article "al" that should have been stripped from words such as "albañil", "alcade", "alfonbra", "almohada", and, last but not least, "álgebra".

[7]Notice, however, that exactly the ability tested here has been developed by Chinese scholars wanting to describe language sounds well before any alphabetic script was introduced into China (Kratochvíl, 1968, pp. 48-49). The pronunciation of a character is described either by comparison with one well known character having the same sound, or by comparison with two characters, one having the same initial consonantal value and the other the same final value (e.g., /lán/ could for instance be described by /luó/ and by /tán/, Kratochvíl, 1968, p. 49). Granted that the subjects used by Read, Zhang, Nie, and Ding (1984) probably spend little time musing about speech sounds, eventually acquiring the tested ability only if confronted with *pinyin*, the procedure just described is nevertheless commonly used in dictionaries.

[8]A somewhat related issue is the monogenesis hypothesis. Gelb (1963) argued that all writing systems may have a common ancestor ; namely, the Sumerian logography. Such an idea cannot be entertained unless the Maya script is classified as a semasiographic forerunner of writing. It seems, however, that there is a growing tendency to consider the Maya script to be highly logographic and even showing phonetization tendencies, though this latter possibility is much more controversial (Kelley, 1976). Even restricted to the Old World, the monogenesis hypothesis is probably doomed to stay highly speculative for ever. Although there is no proof to the contrary, Gelb's arguments for a common origin in the Near Eastern logographic scripts of both the Proto-Indic and the Chinese script are very tenuous at best.

REFERENCES

Ai, J. W. A report on psychological studies of the Chinese language in the past three decades. *Journal of Genetic Psychology*, 1950, *76*, 207-220.

Alegria, J., & Morais, J. Le développement de l'habileté d'analyse phonétique consciente de la parole et l'apprentissage de la lecture. *Archives de Psychologie*, 1979, *183*, 251-270.

Alleton, V. *L'écriture chinoise* (*que sais-je* n° 1374, 2nd ed.). Paris: Presse Universitaire de France, 1976.

Backhouse, A. E. Aspects of graphological structure in Japanese. *Visible Language*, 1984, *18*, 219-228.

Barr, J. Reading a script without vowels. In W. Haas (Ed.), *Writing without letters* (pp. 71-100). Manchester: Manchester University Press, 1976.

Bentin, S., Bargai, N., & Katz, L. Orthographic and phonemic coding for lexical access: Evidence from Hebrew. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 1984, *10*, 353-368.

Berry, J. The making of the alphabets. In E. Siversten (Ed.), *Proceedings of the Eighth International Congres of Linguistics* (pp. 752-764). Oslo: Oslo University Press, 1958. Reprinted in J. A. Fishman (Ed.), *Readings in the sociology of language* (pp. 737-753). The Hague: Mouton, 1958.

Berry, J. 'The making of alphabets' revisited. In J. A. Fishman (Ed.), *Advance in the creation and revision of writing systems* (pp. 3-16). The Hague: Mouton, 1977.

Bertelson, P., Mousty, P., & D'Alimonte, G. A study of Braille reading: 2. Patterns of hand activity in one-handed and two-handed reading. *Quarterly Journal of Experimental Psychology*, 1985, *37A*, 253-256.

Bloomfield, L. *Language*. New York: Henry Holt and Company, 1933.

Bolinger, D. W. Visual morphemes. *Language*, 1946, *22*, 333-340.

Chomsky, N. Phonology and reading. In H. Levin & J. F. Williams (Eds), *Basic studies in reading* (pp. 3-18). New York: Basic Books, 1970.

Chomsky, N., & Halle, M. *The sound pattern of English*. New York: Harper and Row, 1968.

Cohen, M. *La grande invention de l'écriture et son évolution* (3 vols.). Paris : Librairie Klincksieck, 1958.

Coltheart, M., Patterson, K. E., & Marshall, J. C. (Eds.). *Deep dyslexia*. London: Routledge & Kegan Paul, 1980.

Content, A. L'analyse phonétique explicite de la parole et l'acquisition de la lecture. *L'Année Psychologique*, 1984, *84*, 555-572.

Content, A. Le développement de l'habileté d'analyse phonétique de la parole. *L'Année Psychologique*, 1985, *85*, 73-99.

Diringer, D. *The alphabet: A key to the history of mankind* (2 vols.). London: Hutchinson, 1968.

Edgerton, W. F. Ideograms in English writing. *Language*, 1941, *17*, 148-150.

Edgerton, W. F. On the theory of writing. *Journal of Near Eastern Studies*, 1952, *11*, 287-290.

Ehrlich, K., & Rayner, K. Pronoun assignment and semantic integration during reading: Eye movements and immediacy of processing. *Journal of Verbal Learning and Verbal Behavior*, 1983, *22*, 75-87.

Ehrlich, S. F., & Rayner, K. Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior*, 1981, *20*, 641-655.

Février, J. G. *Histoire de l'écriture* (rev. ed.). Paris: Payot, 1959.

Firth, J. R. Sounds and Prosodies. *Transactions of the Philological Society*, 1948, 127-152. Reprinted in E. P. Hamp, F. W. Householder, & R. Austerlitz (Eds.), *Reading in linguistics 2* (pp. 175-191). Chicago: University of Chicago Press, 1966.

Francis, N. W. Linguistics and reading: A commentary on Chapters 1 to 3. In H. Levin & J. P. Williams (Eds.), *Basic studies in reading* (pp. 43-56). New York: Basic Books, 1970.

Frazier, L., & Rayner, K. Making and correcting errors during sentence comprehension: Eye movements in the analysis of structuraly ambiguous sentences.

Cognitive Psychology, 1982, *14*, 178-210.
French, M. A. Observation on the Chinese script and the classification of writing systems. In W. Haas (Ed.), *Writing without letters* (pp. 101-129). Manchester: Manchester University Press, 1976.
Gelb, I. J. *A study of writing* (rev. ed.). Chicago: The University of Chicago Press, 1963.
Gleitman, L. R., & Rozin, P. The structure and acquisition of reading. I: Relations between orthographies and the structure of language. In A. S. Reber & D. L. Scarborough (Eds.), *Toward a psychology of reading: The proceedings of the CUNY conferences* (pp. 1-53). New York: John Wiley & Sons, 1977.
Godwin, C. D. Writing foreign terms in Chinese. *Journal of Chinese Linguistics*, 1979, *7*, 246-266.
Gould, S. J. *Ontogeny and phylogeny*. Cambridge, MA: The Belknap Press of Harvard University Press, 1977.
Haas, W. Writing: The basic options. In W. Haas (Ed.), *Writing without letters* (pp. 131-208). Manchester: Manchester University Press, 1976.
Harden, D. *The phoenicians*. London: Pelican Books, 1971.
Henderson, L. *Orthography and word recognition in reading*. London: Academic Press, 1982.
Henderson, L. (Ed.). *Orthographies and reading*. London : Lawrence Erlbaum, 1984.
Holender, D., & Peereman, R. Differential processing of phonographic and logographic single-digit numbers by the two hemispheres. *Haskins Laboratories: Status Report on Speech Research*, 1985, *SR-81*, 253-290. To appear in G. Deloche & X. Seron (Eds.), *Mathematical disabilities: A cognitive neuropsychological perspective*. Hillsdale, NJ: Lawrence Erlbaum.
Hung, D. L. & Tzeng, O. J. L. Orthographic variations and visual information processing. *Psychological Bulletin*, 1981, *90*, 377-414.
Just, M. A., & Carpenter, P. A. A theory of reading: From eye fixations to comprehension. *Psychological Review*, 1980, *87*, 329-354.
Katz, L., & Feldman, L. B. Relation between pronunciation and recognition of printed words in deep and shallow orthographies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 1983, *9*, 157-166.
Kavanagh, J. F., & Mattingly, I. G. (Eds.). *Language by ear and by eye: The relationship between speech and reading*. Cambridge, MA: The MIT Press, 1972.
Kavanagh, J. F., & Venezky, R. L. (Eds.). *Orthography, reading, and dyslexia*. Baltimore: University Park Press, 1980.
Kelley, D. H. *Deciphering the Maya script*. Austin: The University of Texas Press, 1976.
Kratochvíl, P. *The Chinese language today*. London: Hutchinson University Library, 1968.
Leong, C. K. Hong Kong. In J. Downing (Ed.), *Comparative reading* (pp. 383-402). New York: Collier-Macmillan, 1973.
Leong, C. K., Wong, S., Wong, A., & Hiscock, M. Differential cerebral involvement in perceiving Chinese characters: Levels of processing approach. *Brain and Language*, 1985, *26*, 131-141.
Levitt, J. Spelling-pronunciation in modern French: Its origin and its functional significance. *Linguistics*, 1968, *42*, 19-28.
Levitt, J. The influence of orthography on phonology: A comparative study (English, French, Spanish, Italian, German). *Linguistics*, 1978, *208*, 43-67.
Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. Perception of the speech code. *Psychological Review*, 1967, *74*, 431-461.
Liberman, I. Y. A language-oriented view of reading and its disabilities. *Haskins Laboratories: Status Report on Speech Research*, 1982, *SR-70*, 53-75. To appear in H. Myklebust (Ed.), *Progress in reading disabilities* (Vol. 5.). New York: Grune & Stratton.
Liberman, I. Y., Liberman, A. M., Mattingly, I. G., & Shankweiler, D. P. Orthography and the beginning reader. In J. F. Kavanagh & R. L. Venezky (Eds.), *Orthography, reading, and dyslexia* (pp. 137-153). Baltimore: University Park Press, 1980.

Lotz, J. How language is conveyed by script. In J. F. Kavanagh & I. G. Mattingly (Eds.), *Language by ear and by eye: The relationship between speech and reading* (pp. 125-129). Cambridge, MA: The MIT Press, 1972.

Lyons, J. Phonemic and non-phonemic phonology. *International Journal of American Linguistics*, 1962, *28*, 127-133. Reprinted in E. C. Fudge (Ed.), *Phonology* (pp. 190-199). London: Penguin Books, 1973.

Lyons, J. *Semantics* (Vol. 1). Cambridge: Cambridge University Press, 1977.

Martin, S. E. Nonalphabetic writing systems: Some observations. In J. F. Kavanagh & I. G. Mattingly (Eds.), *Language by ear and by eye: The relationship between speech and reading* (pp. 133-147). Cambridge, MA: The MIT Press, 1972.

Mattingly, I. G. Reading, the linguistic process, and linguistic awareness. In J. F. Kavanagh & I. G. Mattingly (1972). *Language by ear and by eye: The relationship between speech and reading* (pp. 57-80). Cambridge, MA: The MIT Press, 1972.

Mattingly, I. G. The psycholinguistic basis of linguistic awareness. *Haskins Laboratories: Status Report on Speech Research*, 1979, *SR-57*, 155-160.

Mattingly, I. G. Reading, linguistic awareness, and language acquisition. In J. Downing & R. Valtin (Eds.), *Language awareness and learning to read* (pp. 9-25). New York: Springer-Verlag, 1984.

Menninger, K. *Number words and number symbols: A cultural history of numbers*. Cambridge, MA: The MIT Press, 1969.

Miller, R. A. *The Japanese language*. Chicago: the University of Chicago Press, 1967.

Morais, J., Cary, L., Alegria, J., & Bertelson, P. Does awareness of speech as a sequence of phones arise spontaneously? *Cognition*, 1979, *7*, 323-331.

Morton, J., & Sasanuma, S. Lexical access in Japanese. In L. Henderson (Ed.), *Orthographies and reading* (pp. 25-42). London: Lawrence Erlbaum, 1984.

Mousty, P., & Bertelson, P. A study of Braille reading: 1. Reading speed as a function of hand usage and context. *Quarterly Journal of Experimental Psychology*, 1985, *37A*, 217-233.

Navon, D., & Shimron, J. Reading Hebrew: How necessary is the graphemic representation of vowels? In L. Henderson (Ed.), *Orthographies and reading* (pp. 91-102). London: Lawrence Erlbaum, 1984.

O'Neil, W. Our collective phonological illusions: Young and old. In J. F. Kavanagh & I. G. Mattingly (Eds.), *Language by ear and by eye: The relationship between speech and reading* (pp. 111-116). Cambridge, MA: The MIT Press, 1972.

Oommen, C. India. In J. Downing (Ed.), *Comparative reading* (pp. 403-425). New York: Collier-Macmillan, 1973.

Paradis, M., Hagiwara, H., & Hildebrandt, N. *Neurolinguistic aspects of the Japanese writing system*. New York: Academic Press, 1985.

Patterson, K. E., Marshall, J. C., & Coltheart, M. (Eds.). *Surface dyslexia: Neuropsychological and cognitive studies of phonological reading*. Hillsdale, NJ: Lawrence Erlbaum, 1985.

Peereman, R. & Holender, D. Visual-field differences for a number-non-number classification of alphabetic and ideographic stimuli. *Quarterly Journal of Experimental Psychology*, 1985, *36A*, 197-216.

Pollatsek, A., Bolozky, S., Well, A. D., & Rayner, K. Asymmetries in the perceptual span for Israeli readers. *Brain and Language*, 1981, *14*, 174-180.

Pulgram, E. The typologies of writing systems. In W. Haas (Ed.), *Writing without letters* (pp. 1-28). Manchester: Manchester University Press, 1976.

Rabin, C. Spelling reform--Israel 1968. In J. A. Fishman (Ed.), *Advance in the creation and revision of writing systems* (pp. 148-176). The Hague: Mouton, 1977.

Rayner, K. (Ed.). *Eye movements in reading: Perceptual and language processes*. New York: Academic Press, 1983.

Read, C., Zhang, Y. F., Nie, H. Y., & Ding, B. Q. The ability to manipulate speech sounds depends on knowing alphabetic spelling. Paper presented at the *Twenty-third International Congress of Psychology*, Acapulco, Mexico, September 1984.

Rosén, H. B. *Contemporary Hebrew*. The Hague: Mouton, 1977.

Rozin, P., & Gleitman, L. R. The structure and acquisition of reading. II: The

reading process and the acquisition of the alphabetic principle. In A. S. Reber & D. L. Scarborough (Eds.), *Toward a psychology of reading: The proceedings of the CUNY conferences* (pp. 55-141). New York: John Wiley & Sons, 1977.

Sakamoto, T. Reading of Hiragana. In J. F. Kavanagh & R. L. Venezky (Eds.), *Orthography, reading, and dyslexia* (pp. 15-24). Baltimore: University Park Press, 1980.

Sakamoto, T., & Makita, K. Japan. In J. Downing (Ed.), *Comparative reading* (pp. 440-465). New York: Collier-Macmillan, 1973.

Sampson, G. *Writing systems*. London: Hutchinson, 1985.

Sasanuma, S. Acquired dyslexia in Japanese: clinical features and underlying mechanisms. In M. Coltheart, K. E. Patterson, & J. C. Marshall (Eds.), *Deep dyslexia* (pp. 48-90). London: Routledge & Kegan Paul, 1980.

Sasanuma, S. *Surface dyslexia and dysgraphia: How are they manifested in Japanese?* In K. E. Patterson, J. C. Marshall, & M. Coltheart (Eds.), *Surface dyslexia: Neuropsychological and cognitive studies of phonological reading*. Hillsdale, NJ: Lawrence Erlbaum, 1985.

Schane, S. A. *French phonology and morphology*. Cambridge, MA: The MIT Press, 1968.

Seeley, C. (Ed.). Aspects of the Japanese writing system (Special issue). *Visible Language*, 1984a, *18*(3).

Seeley, C. The Japanese script since 1900. *Visible Language*, 1984b, *18*, 267-301.

Skousen, R. English spelling and phonemic representation. *Visible Language*, 1982, *16*, 28-38.

Sun, F., Morita, M., & Stark, L. W. Comparative patterns of reading eye movement in Chinese and English. *Perception and Psychophysics*, 1985, *37*, 502-506.

Sung, M. M. Y. Chinese language and culture: A study of homonyms, lucky words and taboos. *Journal of Chinese Linguistics*, 1979, *7*, 15-28.

Tauli, V. Speech and spelling. In J. A. Fishman (Ed.), *Advance in the creation and revision of writing systems* (pp. 17-35). The Hague: Mouton, 1977.

Taylor, I. The Korean writing system: An alphabet? A syllabary? A logography? In P. A. Kolers, M. E. Wrolstad, & H. Bouma (Eds.), *Processing of visible language 2* (pp. 67-82). New York: Plenum Press, 1980.

Turvey, M. T., Feldman, L. B., & Lukatela, G. The Serbo-Croatian orthography constrains the reader to a phonologically analytic strategy. In L. Henderson (Ed.), *Orthographies and reading* (pp. 81-89). London: Lawrence Erlbaum, 1984.

Twine, N. The adoption of punctuation in Japanese script. *Visible Language*, 1984, *18*, 229-237.

Tzeng, O. J. L., Hung, D. L., & Wang, W. S. Y. Speech recoding in reading Chinese characters. *Journal of Experimental Psychology: Human Learning and Memory*, 1977, *3*, 621-630.

Vachek, J. Some remarks on writing and phonetic transcription. *Acta Linguistica*, 1945-1949, *5*, 86-93. Reprinted in E. P. Hamp, F. W. Householder, & R. Austerlitz (Eds.), *Readings in linguistics 2* (pp. 152-157). Chicago: University of Chicago Press, 1966.

Vachek, J. Two chapters on written English. *Brno Studies in English*, 1959, *1*, 7-38.

Vachek, J. *Written language: General problem and problem of English*. The Hague: Mouton, 1973.

Venezky, R. L. Principles for the design of practical writing systems. *Anthropological Linguistics*, 1970a, *12*, 256-270.

Venezky, R. L. Regularity in reading and spelling. In H. Levin & J. P. Williams (Eds.), *Basic studies in reading* (pp. 30-42). New York: Basic Books, 1970b.

Venezky, R. L. *The structure of English orthography*. The Hague: Mouton, 1970c.

Wang, W. S. Y. The Chinese language. *Scientific American*, February 1973, pp. 50-60.