

The Universal Grammar of Reading

Charles A. Perfetti

University of Pittsburgh

Reading has universal properties that can be seen across the world's writing systems. The most important one is the universal language constraint: All writing systems represent spoken languages, a universal with consequences for reading processes. These consequences are seen most clearly at the broad principle level: the principle that reading universally requires the reader to make links to language at the phonological and morphemic levels. At the same time, the nature of the writing system and the various orthographies that instantiate it do make a difference for important details of the reading process. Drawing on observations and research from Chinese and Korean, I examine these universal and writing-specific aspects of reading. I also consider the implications of the universal language constraint for learning to read.

There are several important aspects of reading that, so I claim, are interrelated by a central fact about literacy: Writing systems encode spoken language. I refer to this claim as the Language Constraint on Writing Systems. To appreciate that this claim has substance, it needs to be contrasted with what is *not* true—that writing systems directly encode meaning. The force of this Language Constraint is twofold:

1. It blocks any attempt to suggest that reading is a parallel language system. If it were, then writing systems, at least some of them, should directly encode meanings, the way spoken language does.
2. It means that learning how to read must involve learning how one's writing system goes about encoding one's spoken language.

In what follows, I attempt to explain the importance of these two implications. But because this argument can be misconstrued to imply a too-simple conclusion about reading, I then try to draw some distinctions that give matters their due com-

plexity. I begin by describing possibilities for a Universal Grammar of Reading, an umbrella for a set of proposals that derive from the Language Constraint.

A UNIVERSAL GRAMMAR OF READING

To be sure, Universal Grammar is a rather grand phrasing for the simple idea I have in mind—the universal part is not so much a grammar as a principle—so, I apologize for being grandiose in the hope that grandiosity will be a forgiven feature of a Presidential Address. The basic idea is expressed as follows:

(1) Reading : Writing System → Language

Proposition 1 asserts that reading is jointly defined by a language and by the writing system that encodes the language. The language part is to be taken seriously and cannot be identified simply with strings of spoken phonemes.

(2) Language → Grammar + Phonology + Pragmatics
 Grammar → Syntax + Morphology
 Morphology → Lexical Roots + Inflections
 Lexical Roots → Syntactic Categories + Meaning

Proposition 2 asserts, conventionally with standard linguistic formulations, that language is an abstract system that includes well-structured subcomponents. The interpretation of language tokens—a language's spoken and written sentences—is jointly determined by these components. The components of the system of most importance for reading are phonology and grammar, in particular the morphological subcomponent of grammar. The pragmatic principles that powerfully affect language use are important in reading to an extent that is not completely delineated and are beyond the scope of my argument. They are not included in the following analyses.

(3) Writing System → Mapping Principles + Orthography
 Mapping Principles → Graphic Units + Language Levels
 Orthography → Mapping Details

The third proposition is that writing systems can be understood at two levels, a higher level of mapping principles and a lower level of spelling or orthographic constraints. The mapping principles are broad enough to include many different languages. The orthographic constraints are definitionally language specific, although they may be very similar across closely related languages.

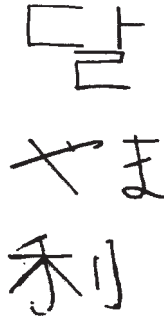


FIGURE 1 Three scripts representing three different writing systems—alphabetic, syllabic, and logographic, from top to bottom. Korean (at top) is alphabetic, with the letters arranged in a square. Thus a square encodes a syllable of from two to four letters. Japanese Kana (middle) directly represents a spoken syllable. The Chinese character at the bottom represents a word, which happens also to be a syllable.

The heart of these ideas is simple: Reading is embedded in two interrelated systems: the Language System and the Writing System. The relation between the first and the second is variable but persistent. There are no writing systems currently in use that bypass language to erect an independent system of signs.

Writing systems work in three ways, according to most systems of classification (e.g., Gelb, 1952): Alphabetic, Syllabic, and Logographic.¹ It is a telling point that all three of these systems can be seen in scripts (which are not the same as writing systems) that appear foreign to an eye used to the alphabets that encode European languages, including English. Figure 1 illustrates this point. Presented are three examples that are visually distinct from those of English and all European and Middle Eastern systems. Furthermore, the three are suggestively similar to each other, when viewed by an uninformed observer. Indeed, they are more similar to each other in visual form than any of them is to English, Dutch, or Hebrew.

In fact, Figure 1 presents not one, but three different writing systems, the entire set that is traditionally defined by writing scholars. Each graph corresponds to a single syllable. The middle row shows the Kana syllabary of Japanese with the logographic Chinese at the bottom. But it is the top panel that is especially interesting because it shows an alphabetic system, Korean. The Korean graph contains three letters arranged from left to right, top to bottom. (Other arrangements also occur in Korean.) The lesson for Western eyes is that you can't tell an alphabet by its *ps* and *qs*.

¹I follow the standard use of *logographic* to refer to Chinese, the only current example of such a system, only to simplify the argument. In fact, Chinese may be better classified as morphemic (Leong, 1973) or, better still, a morphosyllabic system (DeFrancis, 1989), a classification that more directly fixes its linkage to spoken language.

Korean: An Alphabetic System Plus a Moral

These examples help us realize that the principles of the writing system are distinct from the visual appearance of the script. The Korean alphabet (*hangul*) is especially interesting because of its origins. Korea as well as Japan were prolific borrowers of things Chinese, and they made do with borrowed Chinese characters for hundreds of years. The characters were not ideal, however, because the native Korean language is unrelated to Chinese. (The full Korean language contains many Chinese borrowings.) This meant that the characters were mapped with considerable complexity—onto Chinese meanings and sounds, onto Korean meanings and sounds, and onto Korean–Chinese combinations. For example, many characters retained their connection to a Chinese syllable pronunciation that was contained within a Korean word. Others mapped the Chinese meaning of the character onto the Korean equivalent. This complexity led to a less-than-ideal system, although one sufficient to serve an educational elite willing and able to invest the effort to learn it.

In the first half of the 15th century, one of the most remarkable of all events in the history of autocratic governments occurred when King Sejong invented an alphabet for Korean and mandated its universal adoption. The democratic impulses of this monarch can be seen in this translation of a 1434 edict:

Let everyone, in the capital and out, exert themselves in the arts of teaching and instruction . . . let all of them search everywhere for men of learning and sophistication, without regard to whether they are of noble birth or mean, earnestly encouraging them and urging them to teach people to read, even women and girls. (as cited in DeFrancis, 1989, p. 189)

It is interesting to consider how this egalitarian appeal anticipated a similar one from Rudolf Flesch in the United States some 500 years later. In his famous 1955 book, *Why Johnny Can't Read*, Flesch said, “There is a connection between phonics and democracy—a fundamental connection. Equal opportunity for all is one of the inalienable rights, and the word method interferes with that right” (p. 130).

Although Rudolf Flesch’s book received a lot of favorable popular press in the United States, Flesch was branded as an extremist in the educational establishment, and his arguments had only modest impact on typical reading practice. As a monarch, King Sejong’s exhortations did not suffer this fate. Instead, by the proclamation of 1446, Korea suddenly, without the benefit of writing system evolution, had an alphabet:

The sounds of our country’s language are different from those of the Middle Kingdom and are not confluent with the sounds of our characters. Therefore, among the ignorant people, there have been many who, having something they want to put into words, have in the end been unable to express their feelings. I have been distressed be-

cause of this, and have newly designed 28 letters, which I wish to have everyone practice at their ease and make convenient for their daily use. (as cited in DeFrancis, 1989, pp. 189–190)

The Korean King and the American Democrat indeed make a compelling combination across the centuries.

Is Chinese an Exception to the Universal Grammar of Reading?

Leaving Korean for now,² I want to consider the case of Chinese because Chinese has been taken by some to stand in opposition to the principle that writing systems encode languages. The popular form of this misconception is that Chinese writing is pictographic, mapping referents and concepts directly. The more sophisticated form of the misconception is that Chinese, although not mainly pictographic, picks out the meaning level of the language to the exclusion of the phonology. This view treats Chinese as a morpheme-based system.

The first misconception is easily put to rest by the observation that only about 1% or 2% of currently used Chinese characters have identifiable pictographic content (DeFrancis, 1989). Figure 2 illustrates the evolution of the form of the character in the clear direction of abstraction and away from pictures. In Figure 2, it is not possible to discern a pictured object in either the regular or the simplified character, no matter how pictographic the original appeared in its discovery in Shang dynasty oracle bones (1000 or more BC). Lest anyone suppose that the abstraction was a recent development, the modern character shown dates to the Han dynasty of the early third century.

The second belief, that Chinese is morphemic, is not really a misconception but rather an incomplete and, therefore, slightly misleading description. The characters do represent morphemes, but they also represent syllables. Thus, a character is morphosyllabic, corresponding not to an abstract formless piece of meaning but usually to a spoken Chinese syllable that is also a morpheme. Thus, to reuse the horse example, the character represents not horsiness but the Chinese single syllable

²As real life usually has it, the full story is more complex than the King Sejong story implies. Korean has some features that create problems that are absent in most languages that are encoded in an alphabet. In particular, Korean (like Chinese) has much homophony at the syllable level. But the Korean alphabet encodes units of Korean syllables within squares (*kulja*) that were probably a carryover from the characters, which are also constant shape syllable units. This means a given Korean syllable graph corresponds to more than one Korean morpheme, and the reliable mapping that occurs at the level of orthography to phonology disappears at the meaning level. Modern Korean indeed deviates from consistent spelling–sound mapping allowing less meaning ambiguity.



FIGURE 2 The rapid loss of pictographic content in Chinese for the character for horse, *ma*³. The leftmost example is from the Shang dynasty. It is followed by examples from successive points in time—from the Great Seal, the Small Seal, and the Clerical. The two rightmost characters are the current regular and simplified characters.

ble word *ma*³ that means horse. This simple fact means that a Chinese character can be read to correspond to a meaning, to a spoken word, or both. Because Chinese does not have graphic elements that correspond to phonemes, it is not alphabetic. But the writing unit does correspond to a meaning-bearing spoken language unit—the syllable. Thus it maps language, as do all writing systems.

Beyond this basic fact, Chinese becomes more interesting. The bulk of its characters are not like horse (*ma*³), which is a simple character not subject to decomposition. Rather, most characters are compounds that combine two or more constituents (components) that can often stand alone as a character to represent a syllable–morpheme. One kind that corresponds to a popular misconception of how Chinese works in the general case is semantic compounding. For example, the character that means *sun* (日, *ri*⁴) combines with the character that means *moon* (月, *yue*⁴) to make a compound that can be understood as *brightness* (明, *ming*²). Note that there are no shared pronunciations among these three characters. Although it is intriguing in its potential for semantic productivity, semantic compounding of this kind is actually rare.

More common is the kind of phonetic–semantic compounding illustrated in Figures 3 and 4. In these phonetic compounds, one character exposes a clue to its meaning and the other (the phonetic) exposes a clue to its pronunciation. In Figure 3, notice the first character is again *ri*⁴ (*sun*), but this time it combines with the character for green (*qing*¹), which donates its pronunciation to the compound as a whole. Thus the compound is also pronounced *qing*² and means sunshine.

What a lovely system Chinese would be if it worked this way generally. A reader could deduce the meaning and the pronunciation of a compound character, provided he or she knew the component pieces. Alas, this is not to be. Although most characters, over 90%, are compounds that contain a potential phonetic component, in most cases the phonetic component does not give a full mapping to the correct syllabic pronunciation. Sometimes the component and the character share

³The number that follows a syllable represents its tone, one of four pitch contours on Mandarin vowels.

ri4	sun	日
qing1	green	青
<hr/>		
qing2	sunshine	晴

FIGURE 3 An example of valid phonetic compounding. The character for *ri4* (sun) combines with the character for *qing1* (green). The resulting compound *qing2* (sunshine) carries the same pronunciation as one of the components.



JING1(frighten)



LIANG2(cool)



JING1(crystal)

FIGURE 4 Examples of compounds with invalid phonetics. The first two compounds share a phonetic radical but do not share pronunciation. The bottom compound one shares no components with the first but is identical in pronunciation.

a phoneme or two, other times nothing at all. On average, the potential phonetic part of the compound is more likely to have a pronunciation different from the character as a whole as it is to match it, even disregarding tone.

Figure 4 shows an example of this kind of invalid phonetic. Here we see characters that have the same pronunciation, but they do not share a component that provides that pronunciation. So the top and bottom characters are pronounced *jing1*

without any graphic component that indicates this shared pronunciation. The middle one shares a phonetic with the top one, but its pronunciation is different. So the phonetics in Chinese are not always helpful for pronunciation; in fact, they are helpful less than half the time.

I suppose one might argue that Chinese is about as good as English in this respect—*chore* and *choir* share phonetics but not pronunciations. Examples of the variability of English spelling–pronunciation mappings are stock-in-trade for some opponents of phonics teaching, as well as the traditional call-to-arms for spelling reformers. The parallel is quite superficial, however, because although letters can have variable mappings, the mappings they have are systematic and constrained. *Ch* can be /c/ or /k/ but it cannot be something else. Furthermore, the mappings in English are increasingly constrained for units larger than the individual letter and are especially reliable at the rime unit (Treiman, Mullennix, Bijeljac-Babic, & Richmond-Welty, 1995).

This brings me to another aspect of the Universal Grammar of Reading. When the writing system's orthography diverges from its phonology, it does so in a way that is helpful for the reader. In particular, the distribution of this divergence is unequal: The divergence of orthography from pronunciation is less for uncommon words than it is for common words. This state of affairs is familiar in the case of English. So-called irregular spellings are more common for high-frequency English words than for low-frequency English words. Chinese shows the same thing. A divergence between the pronunciation of a character and that of its phonetic component is more common for high-frequency characters than for low-frequency characters. This relation is shown in Figure 5, where the concept of predictable pronunciation is termed *validity*, the extent to which the meaning or pronunciation of the character is predictable by one of its components. Thus, validity refers both to whether the character has the same pronunciation as one of its components and to whether it has a meaning related to the meaning of one of its components. As can be seen in Figure 5, both phonetic validity and semantic validity increase as frequency decreases (see Perfetti, Zhang, & Berent, 1992).

Observing this form–frequency relation in two very different systems suggests a generalization, or to throw caution to the wind, a universal: *Across writing systems, orthographies distribute their divergence from phonology in a way that minimizes the pain to the reader.* The particular way this happens has an intriguing parallel with rule-based processes in language. For example, the English past-tense inflection system has both regular (*knit, knitted*) and irregular (*sit, sat*) components. But the regular system dominates for low-frequency words, and indeed the introduction of new words, as in foreign borrowings, mandates the use of the regular system (Berent, Pinker, & Shimron, 1999). For example, although *mouse* has an irregular plural (*mice*), if Micky and Minnie Mouse have family, we refer to them collectively as “the Mouses” not “the Mice.” To draw a parallel between the Grammar of Writing Systems and the Grammar of Morphology, it ap-

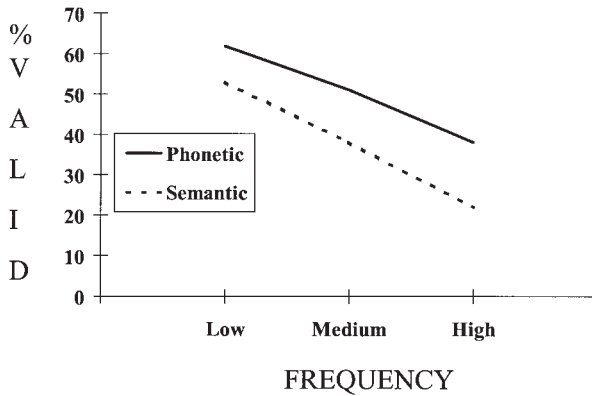


FIGURE 5 The relation between character frequency and validity of semantic and phonetic components. With decreasing character frequency, a character's components give better clues to both pronunciation and meaning.

pears that it is the default rule of Writing Systems, over their evolutionary development, to represent pronunciations. Divergence from this rule—an exception—is permitted more freely for frequently experienced forms than for less frequent forms, honoring a kind of orthographic default across writing systems.

The Universal Phonological Principle

Beyond the logic of writing systems, the facts about how they are read also are important. Not only does Chinese embrace the principle that graphic units represent pronunciation, the research suggests a use of this pronunciation mapping during reading for meaning. The research program that my colleagues and I have carried out comparing Chinese and English has produced ample convergent evidence that reading for meaning in Chinese automatically involves the activation of phonology. Without going into details here, I can highlight our conclusions, as based on tasks of meaning and pronunciation judgment, word naming, lexical decisions, brief-exposure word identification, and Stroop color naming.

For example, in Stroop color naming, we find interference when participants try to name the color blue when the blue color is contained in the print for the Chinese word for red (Spinks, Liu, Perfetti, & Tan, 2000). This much is standard color name interference. But the interesting result is what happens when, instead of *red* (hong2), the word is a homophone of *red* (hong2), a word referring not to color at all but meaning, roughly, *broad*. In the interference condition, the color of the ink—the color to be named—is blue. The word for *red* (hong2) should interfere

with naming the color blue based on the standard Stroop effect—and it does. But interference is also found when the word for *broad* (hong2) is presented. Notice this effect is purely one of pronunciation. The two characters for hong2 share no graphic elements or any meaning. This interference demonstrates clearly that ignoring a character's pronunciation is difficult, just as is ignoring its meaning. The activation of phonology cannot easily be suppressed, as we have found in other tasks as well (e.g., Perfetti & Zhang, 1995).

A typical task from our research is to present readers with pairs of characters and have them decide whether they are related in meaning and pronunciation. In the meaning task, participants viewed one character after another, deciding as quickly as possible whether the two words were related in meaning. On some trials, the two characters were homophones, sharing neither visual form nor meaning but sharing pronunciation. Perfetti and Zhang (1995) found that when the two words were homophones unrelated in meaning, participants showed an interference effect relative to a control condition in which the words were not homophones. This interference result, well replicated in other studies, means that the pronunciation of a character is activated even when the reader's task is to evaluate its meaning. We have recently obtained some neurocognitive evidence that aligns with what we have found in reaction-time experiments. We carried out these experimental tasks while recording event-related potentials (ERPs) from scalp electrodes. At about 250 ms after the onset of the second character, we found that ERP-measured brain activity was affected by whether the word shares pronunciation with the first word. (This homophone effect is a positive-going wave that peaks around 250 ms.) An independent effect of visual similarity is observed 50 ms prior to this homophone effect (Liu & Perfetti, 2002). Thus, we see ERP evidence for two form effects—one of orthographic form and one of phonological form, in that order. Both of these form effects precede meaning effects observed in these tasks.

For the English version of this task, comparative ERP data show some very interesting differences from Chinese. The difference is not in phonology so much as orthography. When American English readers make decisions about whether two words are related in meaning or pronunciation, ERP records show a very early signal associated with similarity of spelling, especially when their task is to decide about pronunciation. It is clear that the brain knows how an alphabetic system works. To put it nontechnically, in these two-word decisions, the alphabetic brain is ready for pronunciation similarity when similar spellings are detected.

We are beginning to fill in bits of the picture about how the brain reads Chinese through collaborative research headed by Li Hai Tan (see also Chee, Tan, & Theil, 1999.) These experiments have shown similarities and differences in comparison with English. The most striking similarity is the activation of left hemisphere frontal regions in meaning tasks, similar to results in English posi-

tron emission tomography studies (Fiez & Petersen, 1998). Especially interesting is that such results are obtained for single-character words as well as two-character words. Previous work on Chinese led to the hypothesis that Chinese single characters were processed nonlinguistically by the right hemisphere and only multiple characters would activate left hemisphere language-processing areas. The results of Tan et al. (2000) demonstrate that even single characters are linguistic objects as far as the brain is concerned. At the same time, there is evidence across a number of our experiments for activation in some areas not seen in English reading, especially the left–middle frontal gyrus (Brodmann area 9), an area that may be associated with spatial and verbal working memory. It is possible that there is a neural basis for the assumption that Chinese indeed involves a visual–verbal process not seen in linear alphabetic writing and perhaps reflecting spatial analysis of character components (Tan et al., 2000, 2001). However, there still is much to be resolved on this question.

If we take a step back to see the big picture over a range of experimental tasks, we come to the clear conclusion that Chinese readers activate phonology even when they read for meaning. However, this does not mean that the writing system exerts no influence on reading. On the contrary, it is clear that the writing system makes a difference in a number of ways, as summarized in Table 1.

One difference is that, although Chinese and English reading both involve automatic activation of phonological forms, English, as an alphabetic system, allows this to occur sublexically in what can be termed *cascade* style. That is, the activation of phonemes based on graphemes accumulates rapidly, cascading to word identification. The activation of a higher unit does not await complete processing of a lower unit but begins immediately even as some graphemes are only partly processed.

In Chinese, the process is different. The activation of phonology awaits a threshold level of graphic recognition before firing. Sublexical phonology (based on a character’s phonetic component) can be observed under conditions that require pronunciation. However, when meaning is the reader’s task, the evidence suggests that the phonology is activated in this graphic-threshold style—the pro-

TABLE 1
Some Comparisons Between English and Chinese

<i>Alphabetic (English)</i>	<i>Logographic (Chinese)</i>
1. Phonology activated with orthography—threshold style	1. Phonology activated with orthography—threshold style
2. Sublexical units: proper parts	2. Sublexical units: wholes are parts
3. Phonology can be “pre-lexical”	3. “Pre-lexical” is not a coherent concept
4. Phonology can “mediate” meaning—but phonological coherence more apt	4. “Mediation” is a dubious concept—phonological diffusion more apt

nunciation of the character as a whole (Zhang, Perfetti, & Yang, 1999). A second, related point is that the phonology of English can be prelexical in that the grapheme–phoneme connections drive phonology from the first moments of visual word processing (Perfetti & Bell, 1991). Chinese does not have prelexical phonology. The phonology that corresponds to a component is syllable-size and morphemic (i.e., the component is also a word). Thus its effect can be considered not prelexical but lexical. In Chinese the part is also a whole.

Another distinction to note is that phonological mediation becomes an awkward concept in Chinese, a fact that follows from the differences I just noted. In fact, the pervasive number of homophones in Chinese—on average, 11 characters share a given syllable/morpheme—makes phonological mediation not very helpful. A pronunciation will nearly always be ambiguous in Chinese. We actually have found mediation effects when the number of homophones for a character is small (Tan & Perfetti, 1997). But the more general point is that phonology provides some constraint on identification, a way of talking that will apply as well to English as to Chinese. Indeed, I think the idea of phonological mediation, as it has been traditionally understood, needs to be abandoned and replaced, as Van Orden and Goldinger (1994) also suggested, by a process that brings about a convergence of identification based on orthographic, phonological, and morphemic constraints.

The fully detailed picture is beyond the scope of this article, so I repeat the main points: The first is that there is no writing system that is read without phonology. Chinese is not an exception to the universal scope of phonology in reading. The second is that the writing system does make a difference. In phonology, the difference it makes is in the details of the orthography-to-phonology mappings. These details are important for the process of reading and for learning to read. There are other differences involving other aspects of the language system—morphology especially, but also syntax—that must be taken into account for a detailed comparison.

The idea of the Language Constraint on reading entails a research program that seeks to understand languages and writing systems to discover the universal principles and also the linguistic and writing system details that control reading. Studies of Korean that we (D. J. Bolger and I) have undertaken with Dr. Hye Kyung Yoon illustrate this idea (Yoon, Bolger, Kwon, & Perfetti, 1999). These studies suggest that, although English readers are highly sensitive to rime units in learning to read and in some reading tasks, Korean readers are not. Instead, Korean readers are sensitive to the syllable body (onset plus vowel). Thus, for an English reader, the structure of the word *sheep* as *sh+eep* (onset plus rime) is functional in learning and in some tasks of skilled reading (but not all; see Booth & Perfetti, 2002). For the Korean reader, the preferred structure is *shee+p* (body + coda), whether they are reading Korean or English. Because similar preferences are obtained in judgments of spoken words, this difference seems to be at the linguistic

level, where the two languages have different syllable structures, not just the reading level. Thus, these studies have suggested that the language itself can affect the level of grapheme–phoneme mappings used in reading. Korean reading for both children and adults is sensitive to body–coda units rather than to the onset–rime units that have proved functional in English reading acquisition.

In the next section, I consider whether these arguments about language constraints and universals have some implications for learning to read. It turns out that the implications, rather than being merely novel, reinforce some common-sense and research-based conclusions.

IMPLICATIONS FOR LEARNING TO READ

The first thing to notice is that the nature of writing systems lies at the heart of the debate about the nature of reading and, by extension, the ways to teach reading. Although the nature of writing systems usually is submerged to the point of invisibility in this debate, it occasionally emerges with startling clarity. In 1995, a rare example of this occurred in Massachusetts, where the State Board of Education had drafted a text for a whole-language approach to its language arts curriculum. In response to the draft of this text, a group of 40 academics, mainly linguists with a sprinkling of psycholinguists, signed a letter criticizing the state’s infatuation with whole language and expressing shock at the naiveté of the whole-language view of reading. This letter of the 40 linguists became notorious in established reading education circles. Reading educators who had sold their philosophy under linguistic banners—first, reading as a “psycholinguistic guessing game” and the reading as “whole language”—saw the rug pulled out from under them by the linguists whose parentage they had taken for granted for so long. Of course, these linguists were dismissed as the wrong kind of linguists: They studied grammar, and the “East Coast” kind (i.e., formal, generative grammar) at that. Such linguists knew nothing about reading, and they were not expert in pragmatics, which is the branch of linguistics traditionally favored by whole-language followers.

But back to the startling clarity on the nature of writing systems: The gang of 40 included the following sentence in their letter: “The proposed Content Chapter [of the State Board’s document] replaces the common sense view of reading as the decoding of *notated speech* [italics added] with the surprising view of reading as directly ‘constructed meaning.’”

I find this sentence to reflect a persuasive point of reference for defining learning how to read. However, I think the notion of “notated speech” not only trivializes written language but also leads to some misunderstanding. I refer to this view of writing systems as the *notational assumption*, and I come back to it at the end.

What Is It That Is Learned in Learning How to Read?

The idea of a Universal Grammar for Reading applied to learning how to read brings into sharp focus the following problem: How does a child learn how his or her writing system encodes his or her language? I have argued that this is the basic formulation of what it means to learn to read (Perfetti & Zhang, 1996). Defining reading as getting meaning from print, however, tends to be the widely accepted view. But the accepted definition confuses one of the goals of reading with the question of the essential nature of reading. Certainly, *one goal* of reading is to obtain meaning (there are others); however, the goal of something is not the same as its essential nature. (We would not accept a definition of playing football as “scoring more points than the other team,” although we would recognize that as the goal.) Moreover, learning to read is about learning, and the accepted definition fails to identify a learning problem in a very useful way. What does it mean to learn to get meaning? What a child learns is how his or her writing system works—both its basic principles and the details of its orthographic implementation. We know this learning has occurred when the child can identify printed words as words in his or her spoken language in a way consistent with the writing system. For an alphabetic reader, this means being able to read unfamiliar words, and even nonwords, as well as familiar words. For a Chinese reader, this means identifying familiar characters and being able to make informed guesses about the pronunciation or meaning of unfamiliar characters using their compositional principles. (This does not include reading nonwords, which in Chinese is impossible, strictly speaking.) To be sure, much more is learned than how one’s writing system encodes one’s language. But this is the central learning event to which additional literacy learning, for example, comprehension strategies, must be connected.

Postulating a universal grammar of reading allows the following question to be considered: Does the child have any prior knowledge about the form of his or her writing system? In the case of language, the answer to the parallel question is usually *yes*. Biological capabilities that are part of species endowment allow humans to acquire language naturally, and part of this endowment includes implicit knowledge of constraints on the forms of human language.⁴

We are less likely to suggest this kind of implicit prior knowledge in the case of writing systems. Whence would such knowledge come? And if children had it, why would demonstrating phonemic awareness prior to literacy be so difficult? And if knowledge of the form of writing systems were available to the species in the way that language is, why did the invention of writing take so long in the course

⁴The specific implementation of this implicit knowledge continues to be a source of debate. However, even on the claim that acquisition of language shares general pattern learning mechanisms (Elman et al., 1996) rather than language-specific mechanisms (e.g., Pinker, 1994), its rapid and untutored nature stands apart from much of ordinary learning.

of human history? Writing is a relatively recent human invention, and its application is far from universal. These facts suggest that writing is a human ability that was laboriously and incrementally extracted from the general well of human creativity, rather than springing from some specialized knowledge. Nevertheless, there is no reason to think that the broad principle that writing connects to spoken language is beyond the grasp of a preliterate child. Indeed, children capitalize on this principle prior to instruction when they begin to write graphs to represent speech sounds. Invented spellings are the child's initial demonstration of the alphabetic principle.

To return to learning: One part of what it means to learn how your writing system works has become central over 20 years of research on phonemic awareness. The importance of phonemic awareness is so well established that I do not say any more about it, except what I think is the most important fact: Children tend not to be able to express this level of awareness outside of literacy. The acquisition of reading universally both supports and is supported by a learned sensitivity to spoken language units. However, the acquisition of alphabetic literacy specifically entails learning about the phoneme level to a degree that is not general across writing systems. The study of Chinese has been helpful on this point. Readers who learned the pinyin alphabet prior to learning the character system are more likely to show awareness of phonemes than readers who did not learn a pinyin alphabet (Read, Zhang, Nie, & Ding, 1986). Readers of a syllabary system appear to acquire some phonemic sensitivity but at a slower pace, as indicated by studies of Japanese Kana (Mann, 1986). However, the fact that the phoneme level of sensitivity develops at all within a syllabary context may depend on specific features of instruction (present in Kana) that draw attention to phonemes (Alegria & Morais, 1991).

However, *phonemic* awareness is not exactly what needs to be learned for reading to get a start. What needs to be learned is that the printed forms on the page correspond to words in spoken language. In alphabetic writing, the smallest of these printed units correspond roughly to small pieces of meaningless speech. In other systems, the smallest printed word corresponds to a spoken syllable. This is true for Chinese as well as for the Japanese Kana system, despite their differences at the level of writing system design. The Kana system promotes a simple one-to-one mapping between a Kana graph and a spoken Japanese syllable. The multiple syllable structure of Japanese words makes it clear the mapping is form-to-form rather than form-to-meaning. The Chinese system is more complex from a learning point of view because, although the graph to be learned corresponds to a syllable, it also corresponds to a specific morpheme, and the number of homophones in Chinese assures that the syllable itself will not identify the morpheme. Nevertheless, on the broader question of mapping graphic units to spoken language units, both systems pick out a syllable level. Thus, both systems exploit and reinforce the child's sense that the writing system maps onto the spoken language. So syllabic, logographic (morphosyllabic), and alphabetic

systems all expose to the child a universal property of writing: The graphs correspond to spoken language units.

On the face of it, learning graphic–syllable units is easier than learning graphic–phoneme units. Syllabic writing was invented prior to alphabetic writing, presumably because syllables were psychologically more accessible. As Gleitman and Rozin (1977) demonstrated, children who are struggling with English might benefit from learning a syllabary first. This benefit is not because learning about syllable mappings will naturally transfer to learning about grapheme–phoneme mappings but because it allows the child to see that graphic units correspond to spoken language. It is this insight that is the heart of a universal language constraint on writing systems.

Although syllables are more accessible than phonemes, it is useful not to overstate the intrinsic difficulty (as opposed to the implementational difficulty) of alphabetic learning. Mappings between letters and phonemes are not intrinsically hard for children to learn. The Chinese situation is again informative. Prior to instruction on characters, Chinese children learn pinyin, an alphabetic system based on the Roman alphabet. This instruction takes 8 weeks, a period during which children learn the sounds that correspond to the letters and then learn to read Chinese sentences written in pinyin. Following this, character learning proceeds by pairing characters with pinyin spellings, before character reading becomes a stand-alone process. In the absence of data that inform us about the effectiveness of pinyin learning in Chinese children, I take this 8-week period to be a reasonable hypothesis about how long it takes to learn the basics of a highly reliable (“regular”) alphabetic writing system implemented for one-syllable words. The reliability condition is important. The pinyin is a completely regular orthography, a one-to-one mapping of letters and sounds. This is an advantage of an invented system. If you invent a system, you make it regular, just as King Sejong did.

The general point is that alphabetic learning of this kind can proceed quickly in principle. For this to work, the basic principle of alphabetic mapping must be implemented in light of practical considerations; in particular, the alphabetic principle can be applied best to simple cases that reflect the principle’s central reliability over its noncentral deviations. Things can get complex in English, but the primary goal remains to help the child quickly learn the principle that organizes his or her writing system. There is time enough to deal with the difference between *chore* and *choir*.

The complexity of English orthography cannot be an excuse for not teaching the principles and exemplary regularities in the system. The scandalous failure to make sure that children learn how the alphabetic writing system works has been widely noted and addressed in recent national reports (National Institute of Child Health and Human Development, 2000; National Research Council, 1998). A recent review of science of reading—cognitive, developmental, cognitive neurosci-

ence, computational, and instructional research—lays out a broad foundation for instructional principles (Rayner, Foorman, Perfetti, Pesetsky, & Seidenberg, 2001). Perhaps teaching how alphabetic writing systems work and how English orthography works in detail will become a more central goal of beginning reading instruction. Still, any expectation that phonics teaching by itself will reverse low literacy levels very dramatically probably will be disappointed.

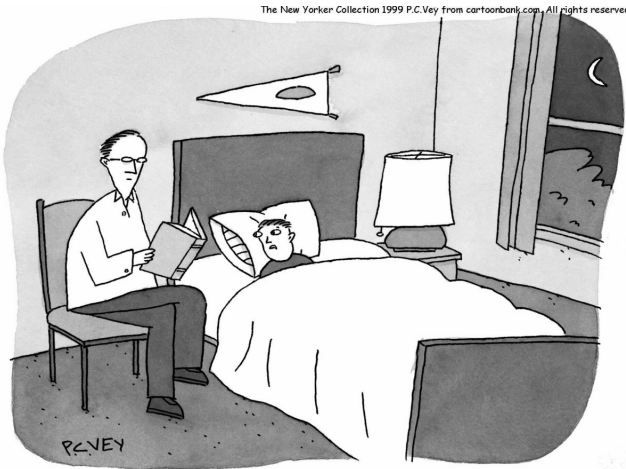
There is so much more to learning to read than the alphabetic principle. Reading shares the fate of other forms of learning: chess, tennis, skiing, fly fishing, or foreign language learning. The principles are as only as good as their specific application to real situations and the effective practice that comes with this application. Reading is not about prolonged acquisition of the alphabetic principle. It is about the accumulation of printed word representations such that a reader can read individual words, using the alphabetic principle as part of the process.

There thus are two classes of learning events in word reading. The child must learn the principle of the writing system and the details of the orthography. Effective instruction and a little insight deliver the first. Experience—lots of experience—with individual word events, reading, and conventional spelling brings the second.

The Elements of Reading Practice

How is experience in reading to come about? Practice is the general answer, but this is too big an issue to be treated here fully. The parts of practice that are related to my main argument are the opportunities to map spoken language to print and then to practice this mapping through reading. For the first, the value of reading to children at home has long been appreciated and has also been recognized in recent national reports. There is much potential value to the child in hearing spoken language associated with printed pages, even if at first the child might not know the source of that language is the squiggles on the page. So indeed, it is sensible to endorse calls to parents to read to their children, even aside from the substantial noncognitive benefits that come from parent–child interactions around books. But beyond being read to, actual reading is necessary.

Early practice at reading—at home—is oral reading, and rightly so. Parents read aloud to children; later, children begin to demonstrate their learning by reading aloud to adults. Why reading education in the first grade often devotes most of its reading time to independent (i.e., silent) reading is a mystery, inscrutable to those who take a more common sense view of the relation between reading and language (see Figure 6). Certainly, from the point of view of what is to be learned in reading, oral reading is what is needed to bring home speech–print connections. Children actually seem to grasp very well the importance of read-



“Shouldn't you be reading that to me out loud or something?”

FIGURE 6 Cartoon representation of the relation between reading and language. *New Yorker* cartoon by P.C. Vey, The New Yorker Collection, 1999, from Cartoonbank.com. All rights reserved.

ing out loud, and they distinguish it from other reading-like (or emergent literacy) activities.⁵

Of course, practice at silent reading must come, and children seem to naturally move to silent reading as they gain skill and wish to mimic what they see adults doing when they read alone. Fluency and comprehension, the twin components of reading beyond the beginning, are nurtured by both oral reading and silent reading.

WRITING AS NOTATION RECONSIDERED

In concluding, I return to a lesson to be learned from Korean and how it applies to the notational assumption that writing is a notation for speech. The thing to recall is the elegant phonemic transparency of the Korean system, as it was envisioned

⁵ If I may be forgiven one child-of-the-author story, my daughter Angela at age 4 had been so immersed in parental reading that she had managed to memorize long passages from stories that were read to her. One night, as I opened up one of her favorites, *Beauty and the Beast*, she told me that she could read the first part. She began with the first word, moved her eyes knowingly from one section to the next, and after some three pages, she had given a flawless word-by-word reading of the beginning of the story. She stopped to say that she had remembered all she could. She said she couldn't go on until she could “read the words like you do.” There is nothing special about this kind of thing; it is commonplace among preliterate children to play at reading while also knowing that reading is something rather specific about what someone does with print.

and implemented by King Sejong. One graph was one phoneme. The graphs were grouped into syllable size chunks called Kulja. But there was a problem. The syllables tend to be ambiguous at the morpheme level (as Chinese is also). So a reader can decode a syllable but not, without context, be sure what morpheme it is. Furthermore, morphology affects the phonology. A word stem that is pronounced one way in its uninflected form is pronounced in a different way in its various inflected forms.

In adaptation to these problems, Korean underwent a change. Instead of spelling the surface pronunciation of all words, spellings for many words began to reflect morphemic structure, which sometimes diverged from phonemic structure. Figure 7 shows an example.

The lexical morpheme /got/ (*flower*; Figure 7, top row) undergoes phonological change as its grammatical case is altered in different syntactic environments, as seen in the second and third rows of Figure 7. In the classical alphabetic system invented by King Sejong, these changes in pronunciation were marked with changes in spelling, following the rule of one letter for each phoneme. In the modern system, the fact that all three morphological variants share the lexical stem is preserved in their spelling. In fact, all three contain the spelling for /gochi/. *Gochi* in the classical spelling was essentially *go+chi*, so one could not see the *got*. Now it is

Spoken form	Meaning	Classical orthography (15C-19C)	Modern orthography (20C-)
/got/	[flower] _{neutral}	꽃	꽃
/gochi/	[flower] _{agent}	꽃치	꽃이
/gotkwa/	[flower] _{with}	꽃과	꽃과

FIGURE 7 An illustration of how Korean has moved from a pure alphabetic to a morphologically sensitive alphabetic system. The lexical morpheme /got/ (*flower*; top row) undergoes phonological change as its case is altered in different syntactic environments, as seen in the second and third rows. In the classical alphabetic system, these changes in pronunciation were marked with changes in spelling. In the modern system, the fact that all three morphological variants share the lexical stem /got/ is preserved in their spelling. Example due to Professor Kwang Oh Yi of Yeungnam University, Korea.

essentially *goch+i*, clearly marking its root morpheme and its inflection. Indeed, even *got* is now spelled *goch*, which becomes the root morpheme spelling across all variations. The instrumental *got+kwa* is now *goch+kwa*.

A movement from a pure alphabetic system to a morphologically sensitive alphabetic system is adaptive for languages like Korean. It reduces the meaning uncertainty associated with a phonological form and allows morphology to be expressed in the writing. This mixed system is superficially similar to what happens in English for derivational pairs like *nation* and *national*. In both cases, there is a compromise on phonological mapping to preserve morphological form. Because the pronunciation alternation patterns in English are predictable within the phonological system, the cost of the variability to alphabetic reading is minimal. The same is true for Korean, although the morphological system carries much more of the predictability.

The Korean example illustrates an important fact about alphabetic writing systems. They are not generally notational systems for speech but notational systems for language, which means the morphology as well as the phonology. That does not mean the notational assumption is wrong. It does mean that we cannot identify spoken language or speech with phoneme strings. When speakers produce a word, they produce a morphological, syntactic, and phonological object. An alphabetic system picks out the phonemic level (part of the phonological specification of the word, which includes more than a string of phonemes) and then makes various adaptations to the morphological level. As the Korean case shows, even a pure phonological system that is the product of rational invention is subject to compromise. Perhaps this news should help educators feel a bit more comfortable with English. Not all of its variation in spelling–pronunciation mapping is systematic in preserving morphology. Indeed, a good deal of it is quite unpredictable without knowledge of language roots (especially Germanic and Latinate) and knowledge of pronunciation changes that occurred long ago. But when we add the fact that English is much more predictable when units larger than graphemes are considered, the obstacles to print–language mappings in English are not insurmountable.

Instruction in learning to read is right to focus on mappings between letters and phonemes. Getting that part roughly right carries the subtleties of morphology along with it fairly readily. Perhaps a child learning to read will make a stab at *jumped* and produce /jump' ed/. But getting to /jump/ is not a big move because that is the spoken form the child has. At least, this is not a big move if the child is getting the central idea—that what the child sees in print maps onto his or her spoken language. Getting this idea across clearly, regardless of the writing system and the complexity of its orthography, is the simple implication of the universal language constraint.

ACKNOWLEDGMENTS

This article is based on the Presidential Address to the Society for the Scientific Study of Reading, Boulder, Colorado, June 2001 and has benefitted from helpful comments by C. K. Leong.

REFERENCES

- Alegria, J., & Morais, J. (1991). Segmental analysis and reading acquisition. In L. Rieben, & C. A. Perfetti (Eds.), *Learning to read: Basic research and its implications* (pp. 135–148). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Berent, I., Pinker, S., & Shimron, J. (1999). Default nominal inflection in Hebrew: Evidence for mental variables. *Cognition*, *72*, 127–138.
- Booth, J. R., & Perfetti, C. A. (2002). Onset and rime structure influences naming but not early word identification in children and adults. *Scientific Studies of Reading*, *6*, 1–23.
- Chee, M., Tan, E., & Theil, T. (1999). Mandarin and English single word processing studies with functional magnetic resonance imaging. *Journal of Neuroscience*, *19*, 3050–3056.
- DeFrancis, J. (1989). *Visible speech: The diverse oneness of writing systems*. Honolulu: University of Hawaii Press.
- Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Re-thinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.
- Fiez, J. A., & Petersen, S. E. (1998). Neural imaging studies of word reading. *Proceedings of the National Academy of Sciences USA*, *94*, 914–921.
- Flesch, R. (1955). *Why Johnny can't read*. New York: Harper & Row.
- Gelb, I. J. (1952). *A study of writing*. Chicago: University of Chicago Press.
- Gleitman, L. R., & Rozin, P. (1977). 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. 1–54). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Leong, C. K. (1973). Reading in Chinese with reference to reading practices in Hong Kong. In J. Downing (Ed.), *Comparative reading: Cross-national studies of behavior and processes in reading and writing* (pp. 383–402). New York: Macmillan.
- Liu, Y., & Perfetti, C. A. (2002). *ERP evidence for the time course of graphic, phonological and semantic information in Chinese meaning and pronunciation decisions*. Manuscript submitted for publication.
- Mann, V. A. (1986). Phonological awareness: The role of reading experience. *Cognition*, *21*, 65–92. National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Publication No. 00–4769). Washington, DC: U.S. Government Printing Office.
- National Research Council. (1998). *Preventing reading difficulties in young children* (C. E. Snow, M. S. Burns, & P. Griffin, Eds.). Washington, DC: National Academy Press.
- Perfetti, C. A., & Bell, L. (1991). Phonemic activation during the first 40 ms of word identification: Evidence from backward masking and masked priming. *Journal of Memory and Language*, *30*, 473–485.
- Perfetti, C. A., & Zhang, S. (1995). Very early phonological activation in Chinese reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 24–33.

- Perfetti, C. A., & Zhang, S. (1996). What it means to learn to read. In M. F. Graves, B. M. Taylor, & P. van den Broek (Eds.), *The first R: Children's right to read* (pp. 37–60). New York: Teachers College Press.
- Perfetti, C. A., Zhang, S., & Berent, I. (1992). Reading in English and Chinese: Evidence for a “universal” phonological principle. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 227–248). Amsterdam: North-Holland.
- Pinker, S. (1994). *The language instinct: How the mind creates language*. New York: Morrow.
- Rayner, K., Foorman, B. R., Perfetti, C. A., Pesetsky, D., & Seidenberg, M. S. (2001). How psychological science informs the teaching of reading. *Psychological Science*, 2(Suppl.2), 31–74.
- Read, C., Zhang, Y., Nie, H., & Ding, B. (1986). The ability to manipulate speech sounds depends on knowing alphabetic reading. *Cognition*, 24, 31–44.
- Spinks J. A., Liu, Y., Perfetti, C. A., & Tan, L. (2000). Reading Chinese characters for meaning: The role of phonological information, *Cognition*, 76, B1–B11.
- Tan, L. H., Liu, H. L., Perfetti, C. A., Spinks, J. A., Fox, P. T., & Gao, J. H. (2001). The neural system underlying Chinese logograph reading. *NeuroImage*, 13, 836–846.
- Tan, L. H., & Perfetti, C. A. (1997). Visual Chinese character recognition: Does phonological information mediate access to meaning? *Journal of Memory and Language*, 37, 41–57.
- Tan, L. H., Spinks, J. A., Gao, J. H., Liu, H. L., Perfetti, C. A., Xiong, J., et al. (2000). Brain activation in the processing of Chinese characters and words: A functional MRI study. *Human Brain Mapping*, 10, 16–27.
- Treiman, R., Mullennix, J., Bijeljac-Babic, R., & Richmond-Welty, E. D. (1995). The special role of rimes in the description, use, and acquisition of English orthography. *Journal of Experimental Psychology: General*, 124, 107–136.
- Van Orden, G. C., & Goldinger, S. D. (1994). The interdependence of form and function in cognitive systems explains perception of printed words. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 1269–1291.
- Yoon, H. K., Bolger, D. J., Kwon, O. S., & Perfetti, C. A. (1999, April). *Grapheme-phoneme processes in English and Korean*. Paper presented at the annual meeting of the Society for the Scientific Study of Reading, Montreal.
- Zhang, S., Perfetti, C. A., & Yang, H. (1999). Whole-word, frequency-general phonology in semantic processing of Chinese characters. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 858–875.

Manuscript received March 25, 2002

Accepted June 20, 2002

Copyright of Scientific Studies of Reading is the property of Lawrence Erlbaum Associates and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.