

Chapter 1

Learning to Spell

What is the problem?

Most of children's learning is generative. They learn not just about specific facts or specific actions, but about how to deal with quite new experiences and new situations. Language is one obvious example, and counting is another. We only hear a limited number of sentences when we learn our first language, but a fluent speaker should be able to say anything he or she wants, using sentences never heard before. We learn to count by getting to know a limited number of counting words, but with this knowledge we should be able to count on and on, far beyond the limits of the specific words that we learned.

Learning to read and spell is much the same. What children learn about reading and spelling words should, in the end, make it possible for them to read and spell reasonably well words that they have never seen in written form before. If children memorized the spellings of lists of words and could write those words correctly, and only those, we could conclude that they had learned how to spell words one by one. But this is not what happens, nor what we would want to happen. We want children to be able to spell most words as a result of learning a limited number of words. This book is about how children manage, or in some cases nearly manage, to do so.

The Connection between Language and Literacy

For many years now the idea that literacy is a language-based activity has dominated research on children's reading and writing. For example, an important collection of papers, published in 1972 by Kavanagh and Mattingly, entitled "Language by eye and by ear" summarized the empirical evidence that had been developed up to that time to support the conception of literacy as a language-based activity. The idea shared by

the authors of these papers, and by most researchers on reading since then, is that the core cognitive processes in reading and writing are linguistic because what we learn in literacy acquisition is a *written language*. Therefore, analysing literacy learning requires understanding what a written language is and how oral language and written language are connected—that is, understanding “the linguistic connection.”

This may seem an easy starting point but there is more than one view of what the relationship between oral and written language is. In fact, this has been a matter of debate for some time; the debate permeates discussions in linguistics, the history of writing, the psychology of literacy and education. In this chapter, we will present two different views of how oral and written language are connected, and explore the teaching and learning implications of these. We will argue that, though these are different views, they are actually quite compatible with each other, and that it is a better approach altogether to develop a theory that integrates the two views.

Two Views of the Relation between Oral and Written Language

The *first* is called the *notational view*. According to Olson (1994), it has been assumed since Aristotle’s time that writing is a graphic device for transcribing speech: “written words are the signs of words spoken” (Aristotle, *De interpretatione*). Although scientific revolutions since Aristotle have changed the ways in which we think about the physical world, this classical view of the connection between oral and written language has not been dismissed and continues to receive the explicit or implicit support of linguists (e.g. Bloomfield, 1933; Mattingly, 1972; Saussure, [1916] 1983), historians (e.g. Diring, 1968; Gelb, 1963; Sampson, 1985), psychologists (e.g. Frith, 1985; Cossu, 1999; Treiman, 1993) and educators (Isaacs, 1930; Montessori, 1918).

Recently, Tolchinsky (2003) detailed this conception by exploring the characteristics of orthographies as notational systems, i.e. as artefacts that enable oral language to be encoded, recorded, transported and reproduced in a systematic way. Adopting definitions proposed by N. Goodman (1976) and Harris (1995), she summarized the features of notational systems in general and showed their value for understanding orthographies. A notational system contains a limited set of elements—letters, in the case of an alphabetic orthography—each with a distinctive form. These elements can be copied and identified in spite of variations

in the way that they are copied by different users. The elements are semantically differentiated (i.e. they refer to different elements of what is represented) and can be structured by specific rules (for example, in English orthography we read and write from left to right, top to bottom). Because of these characteristics, notational systems are powerful tools. With a limited set of letters, we can write all the existing words in a language, and even new ones, invented much later than the orthography.

This notational view treats writing as a second-order system—a system of (graphic) signs for (oral) signs. Thus reading and writing are directly related to and entirely dependent on oral language. This conception of the relationship between oral and written language has consequences for theories of how children learn to read and write, and also for how they should be taught. If orthographies are notations for oral languages, children need to learn how this representation works, i.e. what is represented by the orthography and how. Alphabetic orthographies are those in which letters represent phonemes—even if there is no exact correspondence between letters and phonemes. Other ways of representing language through writing are also possible, for example, by using a different unit of analysis of the sounds that make up words: Japanese orthography uses Kana letters to represent syllables rather than phonemes (see Akita & Hatano, 1999, for a more precise description). Within this notational perspective, children must learn how the orthography represents the language that they speak in order for them to learn how to read and write. The notational view of the relationship between oral and written language is easy to understand and we believe that it is implicitly accepted by most people. However, it is not the only view of the connection between oral and written language.

The *second* approach goes beyond the notational view, and treats writing as *written language*—that is, as a system with its own rules to represent meaning, and not only the sounds of oral language. For some linguists (e.g. Siertsema, 1965; Uldall, 1944), a language system goes beyond the way that it is expressed, either in oral or written form. The sounds that we hear in oral language and the letters that we see on the page are only the surface of the language system. The surface representations express meanings that are part of the deep structure of the language.

Because this approach is less familiar to most people than the notational one, it is useful to start from an analogy to the connection between two different oral languages. A sentence in oral language—for example, “The boy chased the dog”—expresses more than each word taken in isolation. The word order indicates that it was the boy that did the

chasing—it signals the subject–verb–object (S–V–O) structure of the sentence. The principle of word order is used in English to represent the underlying grammar and it is what allows us to “generate an infinite class of sentences” (Chomsky, 1975, p. 41). This abstract system is the basis for learning any language in spite of the differences that exist between languages, and it is also used to learn a written language.

The approach to writing as a written language starts from the idea that the same sort of deep structure is the basis for written language as well as for oral language. Thus written language can use its own resources to represent the meaning relations that exist in the grammar of the language, even if they are not captured in the same way in oral language. For example, we use “s” to mark the plural of nouns and we use “ed” to mark the past tense of regular verbs but in oral language plural words can have the ending sounds /s/ or /z/ and past regular verbs never sound as “ed” at the end. In principle, someone who understands about plurals and the past tense can learn how these meanings are marked in written language even when they are not phonological notations of spoken language.

When children learn an oral language, they learn to give a phonological form, which is arbitrary, to the semantic and syntactic relations that they wish to express. Similarly, when they learn to write, they learn to give a graphic, arbitrary form to the semantic and syntactic relations that they wish to express. Thus, in this view, an orthography does not represent only the surface form of oral language: there are also connections between the deep structure of the language, which represents grammar and morphology, and the way in which a language is written.

This conception of how oral and written language are related may seem highly academic and without any pragmatic consequences for teaching and learning, but it is not so. If written language is only a different expression of the same language system that can be expressed orally, it should be learned through its connection to the abstract language system, not through its connection to oral language. K. Goodman (1982), for example, argued that in reading instruction “So-called ‘linguistic programs’ that emphasize phoneme-grapheme correspondences à la Bloomfield and Fries are still emerging, perhaps five or ten years beyond that point where there was any justification at all” (p. 90). He further proposed that

[A]lphabetic systems don’t simply operate on a letter-sound basis. . . . Sequences of sounds seem to have relationships to sequences of letters, not simply because of the alphabetic principle on which the system was

produced originally, but also because there is a common base underlying both of these. For the user of language, surface oral language and surface written language are related through a common underlying structure. As a language user generates a sentence, his thoughts bring him to a point at which he can apply a set of orthographic rules and write it. (pp. 91–92)

This view of how oral language and written language relate to each other implies that when children begin to read and write they learn a set of rules for expressing and understanding sentences in written form. They learn to produce meaningful sentences in writing, in just the same way as they previously learned to give an oral form to sentences in speech.

The second view of how oral language and written language are related is much less likely than the first to fit people's intuitions, and is more difficult to understand. Yet it is quite easy to find support for it. In written English, as in many other orthographies, we make distinctions in writing which are not marked in the sounds of words.¹ For example, we spell the end sounds /ks/ differently in different words: think of "fox" and "socks," "mix" and "tricks," "tax" and "tracks." If we were simply trying to represent the sounds of oral language, why would we spell these word endings differently? Are these spelling differences illogical and entirely unpredictable? Of course not! These spellings are entirely predictable if we think not only about oral language but also about the connection between oral and written language to an abstract language system that represents grammatical relations. The /ks/ sound at the end of words in English is represented by the letter "x," as in "fox," "mix" and "tax"—except when the word can be decomposed into a stem plus an affix, as in "sock+s," "trick+s" and "track+s."

Linguists, such as Chomsky (1965), have argued that we understand sentences by connecting them to an implicit grammar that represents simpler sentences. He argues that this is what allows us to recognize ambiguities: a sentence that we hear is ambiguous when it can be connected to different underlying simpler sentences. To use one of his examples: the sentence "I had a book stolen" could mean that "I had a book; someone stole it" or "someone stole a book; I asked that person to steal the book." It is easy to make an analogy between analysing an ambiguous sentence by connecting it to different core sentences and analysing a sequence of sounds, such as /ks/ at the end of words, by connecting it to different forms of words: the ending /ks/ can be

¹ In this book, we will use quotation marks when we refer to a word or part of a word and a letter between forward slashes when we refer to sounds.

connected to a plural form, /k/+ “s” or to a singular form, which does not contain the letter “s.”

These two views of the relation between oral language and written language have led to diametrically opposed approaches to the teaching of literacy. The first view has emphasized the need to help children become aware of the sounds in their language so that they understand that letters represent sounds. This is sometimes referred to as “attaining an alphabetic conception of written language” (Read, 1971, 1986; Ferreiro & Teberosky, 1983) or “learning the alphabetic principle” (Byrne, 1998). The second view is associated with the idea that children can learn to express language in writing if they are exposed to it (K. Goodman, 1982)—an idea that formed the basis for the “real books” approach to literacy instruction.

Steps towards a Synthesis

In this book, we will pursue a synthesis of these two positions. It is suggested that orthographies are notational systems and, as such, they enable the encoding, recording and reproduction of oral language. However, oral language and written language are not connected only through their surfaces: they are also connected through their relation to an abstract yet specific (e.g. English, French) language system.

The linguist Jean Pierre Jaffré (1997; adopting a modified version of Vachek’s, 1973, definitions) explained this view by proposing that writing combines two principles: phonographic and semiographic. The phonographic principle

is manifested by correspondences between meaningless units of spoken language (phonemes or syllables) and meaningless units of written language (phonograms or syllabograms). The semiographic principle encompasses the units and their functions in the linguistic elements of written language. These units are determined by the morphological structure of the languages in question . . . and by the way in which the written words are assembled. (p. 9)

The semiographic principle is at work, for example, when we spell all regular past verbs with “ed” at the end, irrespectively of whether the endings of the verbs are pronounced as /t/, as in “kissed,” /d/ as in “killed,” or /id/ as in “wanted.” But the semiographic principle is not restricted to the spelling of words as such: it is much more pervasive because it is this principle that we use when we place spaces between

words. Words are not phonological units but units defined by meaning and grammar. We say, for example, “I went to school yesterday” but write “went to,” placing a space between the verb and the preposition, though we do not pronounce them separately. Words can change phonologically depending on the context in which they are spoken: a child may write “I hat to run” (see Nunes & Bryant, 2006, Figure 1.1, p. 27) not because the child cannot hear the difference between /t/ and /d/ but because the child is transcribing the sounds rather than preserving the identity of the word “had” across different phonological contexts.

Orthographies that rely more on the phonographic principle have been called “transparent,” and in these orthographies it is difficult to appreciate the fact that writing and oral language are connected via the deep structure of the language. However, studies of competent adult readers in Italian, a transparent orthography, show that they use both phonographic and semiographic principles in word recognition (we do not review this literature here but the interested reader is referred to the work by Caramazza and his colleagues, in Italian; see Chialant & Caramazza, 1995, for a synthesis). It is much easier to recognize this underlying connection in less transparent orthographies, such as English, where the representation of morphemes often results in spellings that cannot be predicted from the way that the word sounds (such as “fox” and “socks”) or might even contradict what would be expected from the sounds. “Magician” is spelled as it is and not as “magishon,” though children often do write it that way. The wrong spelling is quite a good representation of the word’s sounds, but the correct spelling works at a different level. Even though letter “c” in “magician” no longer represents the sound /k/ as it does in “magic,” the first part of the word represents the word’s meaning very well. The “-ian” ending is an affix that is called an “agentive”: it signifies someone who does something. So, “magic” plus “-ian” is a very good way of representing in writing the meaning “someone who does magic.”

The connection between morphemes and spelling, though not immediately obvious, can be easily understood. Once the regularity of the spelling of morphemes is recognized, words that seem highly irregular, like “magician” and “confession,” can be seen as regular (magic+ian; confess+ion): regularity in English orthography is based not only on phonology but also on the “visual identity of meaningful parts” (Venezky, 1999)—that is, the use of the same spelling for the same morpheme even when its pronunciation changes.

One should not be tempted to simplify English orthography by thinking that “the visual identity of meaningful parts” is a principle that

overrides the representation of sounds. There are many rules that apply when a suffix is added to a base form, which involve changing the visual form in order to preserve the phonological representation. A common rule is doubling letters when we add to a stem a suffix which starts with “e.” In English, a vowel that is followed by a consonant, and no “e” after the consonant, is pronounced differently from one that is followed by a consonant plus “e”: for example, “hat” and “hate,” “hop” and “hope.” The pattern Vowel plus Consonant plus “e,” which is often called the “split digraph,” is well known to teachers and explicitly taught to children in school. When we add a suffix that starts with “e” to a stem, the spelling pattern would be changed into a split digraph, and the vowel would sound differently. In order to preserve its sound, the consonant is doubled at the end of the stem: consider these examples: “tan”–“tanned”; “plan”–“planner”; “clot”–“clotted”; “log”–“logged”; “pot”–“potted”; “pin”–“pinned”; “big”–“bigger”; “sin”–“sinner”; “cut”–“cutter”; “run”–“runner.” This is certainly an added complexity in English orthography, but still reasonably easy to master when we think of written language as representing sounds and also based on a connection to the deep structure of the language.

Subtleties in the Linguistic Connection: Influences across Languages

Borrowing words

Not everything about the linguistic connection in English orthography is so easy. Languages are—and must be—dynamic: new terms can enter into the vocabulary at any time. Some new words might be invented in the language itself, by creating compound words (e.g. spaceship, search-engine) or by composing a new word with morphemes that are already part of the language (e.g. skyscraper). Other novel words can be borrowed from other languages—and then the matter of how they are spelled has to be considered. English orthography honours etymology (Venezky, 1999). This means that the same sounds are often spelled differently in different words because the words come from different languages. Relatively recent borrowing with rare or proscribed consonant and vowel patterns (e.g. “tsunami,” “Chianti”) might have little effect on English orthography in general. These words could be learned by memory because they are restricted in number. However,

[E]xtensive borrowing over a long period of time—with different retention patterns for spellings and changing spelling-sound relations in the original language—frustrate not only the use of etymology for predicting spelling-sound patterns, but the entire enterprise of orthographic rule making. (Venezky, 1999, p. 8)

So, there is one level of the linguistic connection that does not help much when we want to spell words in English, even if we are aware of the fact that etymology plays a role in English spelling.

Spelling with a borrowed alphabet

It is interesting to pursue the idea of borrowing at this point in the introduction, because of the effect that borrowing the Roman letters has had on English orthography. Latin (and the languages that originated from Latin, such as Italian, Spanish, and Portuguese) and Greek have a small number of vowel sounds, which can therefore be represented with a small number of letters. Borrowing the Roman letters to represent the sounds of the English language creates a problem. If we ignore the issue of variations in pronunciation and think of the kind of English called RP (Received Pronunciation or BBC English), we should be able to distinguish 21 vowel phonemes (O'Connor, 1982, p. 153). It is obviously not possible to represent these 21 sounds by setting them in one-to-one correspondence with 5–7 letter-vowels (A, E, I, O, U, Y and W). So we see that borrowing an alphabet means using it creatively. One way of using it creatively is to invent larger units with more than one letter in order to represent sounds that are different from those represented by a single letter. There is no need to discuss here the variations in pronunciation of vowels across different regions of English speakers—but there is also no reason to argue about the fact that the vowels in “hat” and “hate,” for example, are different, and that representing both words with the letters “h,” “a,” and “t” would be to ignore this difference. Thus English orthography uses the borrowed Roman letters creatively: it uses digraphs—that is, two letters—to represent one vowel phonemes. The importance of some digraphs is immediately seen—such as the split digraphs “a-e,” “o-e,” “i-e” and “u-e”—but even these are still not sufficient to get 21 vowel sounds represented. Others are still needed, to differentiate, for example, “did” and “deed,” and “lock” and “look.” But it could be argued that some are not necessary, as they overlap in function with other digraphs—“hope” and “boat” do not

have to be distinguished—and the use of extra digraphs for the same function here does make things more difficult for learners.

One of the issues that will concern us in this book is whether children realize quite soon that it does not work to spell vowel sounds on a one-to-one basis by using vowel letters because there are not enough letters to go around for all the sounds. We will analyse their resources for dealing with the shortage of vowel letters before they master conventional spellings.

We now turn to another aspect of English orthography, which has considerable impact on how words are spelled, but is not really about the connection between oral and written language, direct or indirectly, via the surface or the deep structure.

Form and Function in Written Language

In many written languages, including English, some differences in spelling are a matter of form rather than of function. In English, for example, words can end but cannot start with “ss.” The pattern “ss” is actually the preferred spelling for one-morpheme words ending in the sound /s/: so “ss” is a common spelling for the /s/ sound but it is illegal at the beginning of words.

Similarly, “ck” is used at the end but not at the beginning of one-morpheme words. The use of “ck” instead of “k” at the end of words can be predicted by a relatively simple rule, which is also positional: “ck” is used after words that have only one vowel letter in the rime before the /k/ (e.g. “brick,” “lock,” “lack,” “truck”) and “k” is used after words that have more letters in the rime before the /k/ (e.g. “mink,” “book,” “leak,” “park,” “fork”).

Another positional rule concerns the use of “ay” and “ai”: “ay” is used at the end of stems and can also be used before vowels in the middle of a stem; “ai” is (with very few exceptions) only used in the middle and before consonants (e.g. contrast “may” and “main”; “clay” and “claim”).

We think of these positional restrictions in spelling as a matter of “form” only, because they are not obviously related to a linguistic function, either phonological or morphological, even if they tend to apply to morpheme boundaries. This is an intriguing set of spelling patterns, which we hypothesize could be connected to a learning process different from those that apply when form and function go together.

Conclusion

We started out this introduction by posing the problem of reading and spelling as a generative process. It must be generative in order for us to be able to read and spell words that we have never seen before; but the words that we spell cannot be generated in a whimsical way, without any rules, because others should be able to read the words whose spelling we created.

We argued that linguistic processes are at the basis of this generative process in two ways. A prominent linguistic aspect of spelling is that orthographies represent oral language. Thus the letters that we use to represent a new word which we have not learned must guide the readers to how the word should be pronounced. This process is not a simple one-letter-for-one-sound representation; among other reasons for the lack of one-to-one correspondence is the considerable shortage of vowel letters in English. However, English orthography also uses the semio-graphic principle, in Jaffré's terminology, or the principle that Venezky calls "the visual identity of meaningful parts": if the sounds of a stem change when a suffix is added to it in, for example, pairs such as "magic"—"magician" and "heal"—"health," the stem conserves its visual identity in spite of the change in sounds. The same is true in endings such as "kissed," "opened" and "wanted"—three words that have different end sounds, /t/, /d/ and /id/, respectively, are all spelled in the same way, with "ed," a unit of meaning that marks the regular past tense of verbs.

Finally, we also suggested that the linguistic connection does not tell the whole story. There are in English many restrictions on the use of letters by position, and when there is more than one option for spelling, as in "ay" or "ai," positional rules might be the crucial factor in indicating the correct spelling.

Research on children's learning of word reading and spelling, as well as teaching, has focused to a large extent on the connection between the surface of oral and written language. In this book, we argue that this focus on letter-sound relations is necessary for an understanding of how children learn to read and spell but not sufficient. Most children master simple letter-sound relations in word reading and spelling within one or two years from the beginning of their reading instruction but still have a lot to learn to become good readers and spellers.

The aim of this book is to go beyond studying how children learn simple letter-sound correspondences, in order to consider how children

learn how to use more than one letter to represent a single phoneme and to use semiographs that do not correspond to phonological units. We will try to answer question like: Do children easily see that sometimes two or more letters are used to represent a single sound? Do they seem to learn form rules, such as no doublets at the beginning of words, independently of function rules, such as the need to double consonants when “ed” is added to a verb to form the past tense? When there are different options for spelling the same sound, such as “x” and “ks” at the end of words, do they learn each word separately or do they learn a rule that helps them choose the right spelling? How easy is to teach children about semiographic rules and what are the best ways of teaching these? Although all these are very important issues when children are learning to read and spell, there are not many sources of answers to these question. We provide some answers and invite you to explore them with us.