English Orthography and Reading

RYAN T. MILLER

Framing the Issue

One of the most fundamental aspects of learning to read is understanding how printed text relates to spoken language (Perfetti, 2003). When learning to read in English, a learner must view printed letters (graphemes), decode their sounds, and combine those sounds together to form words. For example, to read the word *cat*, a beginning reader must understand that the grapheme *c* makes a [k] sound, the grapheme *a* makes an [æ] sound, and the grapheme *t* makes a [t] sound, before combining them into [kæt], a word which they already know the meaning of orally. This process of decoding graphic forms into phonological forms is a key component of word recognition, which itself is a key component of learning to read. The specific patterns of correspondences between the graphic and phonological forms are the *orthography* of a language. Each language has its own unique orthography. Thus, all learners who are learning to read in English, no matter their first language background, need to develop their knowledge of the orthography of English.

The *writing system* of a language is a related concept, but is distinct from its orthography. A language's writing system defines the linguistic unit that is represented by the graphemes of a language. There are three main types of writing systems: alphabetic, syllabic, and morphographic (Coulmas, 2003). In alphabetic languages, graphemes represent phonemes or individual sounds. English is an *alphabetic* language, like many other European languages such as Spanish, German, French, and Italian, but also languages such as Arabic, Hebrew, and Korean *hangul*. On the other hand, in a *syllabic* writing system (such as Japanese *kana* or Cherokee), each grapheme represents a syllable; for example, the sounds *ba*, *bi*, *bu*, *be*, and *bo* would each be represented by single graphemes. In a *morphographic* writing system (such as Chinese, Japanese *kanji*, or Korean *hanja*), each grapheme represents a morpheme or a unit of meaning.

Within each system type, variation exists in the specific details about the correspondences between graphic symbols and language. Orthography refers to this language-specific variation. Lastly, *script* refers to the specific symbols that are

The TESOL Encyclopedia of English Language Teaching.
Edited by John I. Liontas.

DOI: 10.1002/9781118784235.eelt0461

used to write a language. Thus, English and Chinese have different writing systems, English and Russian share a writing system (alphabet) but not a script (English uses the Roman script while Russian uses the Cyrillic script), and English and Spanish share a writing system (alphabet) and script (Roman script) but differ in their orthographies (i.e., the specific correspondences between graphemes and sounds).

Making the Case

In the initial stages of learning to read in English, a learner needs to first understand the general mapping principle of English; that is, that, essentially, each letter represents a distinct sound or phoneme. For learners whose L1 is alphabetic, this will generally not be a difficult concept to grasp, while it may be more difficult for those whose L1 is syllabic or morphographic. For example, although the script used in Korean, *hangul*, is different from English, it is nonetheless alphabetic and follows the same general mapping principle as English, with each grapheme representing a phoneme, so the alphabetic principle will be relatively easy for learners with previous literacy in Korean to understand.

This is in contrast to learners with previous literacy experience in a morphographic writing system, such as Chinese. In a morphographic language, readers have learned to associate graphemes to a greater extent with morphemes or meanings, rather than phonemes. Readers who became literate first in a morphographic language also tend to rely more on visual cues (e.g., radicals located within characters) to identify word meanings, while readers of alphabetic languages tend to rely more on phonological information.

These differences in visual processing of words in a reader's first-acquired literacy become the basis on which they read subsequent languages as well. Research has found that L2 readers use the orthographic processing strategies of their L1 or first literacy when reading in an L2 or subsequent literacy (Akamatsu, 2003), and, thus, English learners' reading will vary systematically according to their L1 background, provided that they have literacy experience in that language. For example, Wang, Koda, and Perfetti (2003) found that in an L2 English semantic category judgment task, L1 Korean (alphabetic) readers made more false positive errors on items that were homophones of target words (showing greater reliance on phonological information), while L1 Chinese (morphographic) readers did not show such an effect, and instead showed more false positive errors on items that were visually similar to target words (showing greater reliance on visual information).

In addition to differences in reading that result from variation among writing systems, there is also variation among alphabetic languages in the degree to which graphemes correspond to phonemes. The continuum of this variation is referred to as *orthographic depth*, and the orthographic depth hypothesis predicts how orthographic depth influences reading processes (Katz & Frost, 1992). Languages with very shallow orthographies (e.g., Turkish, Serbo-Croatian, Korean *hangul*) have highly regular, one-to-one grapheme–phoneme correspondence. This facilitates

decoding of written words and encourages readers to analyze words phonetically, letter-by-letter. Slightly less shallow (though still on the shallow end of the scale) are the orthographies of Spanish, Italian, and Greek. Even less shallow (i.e., more irregular grapheme–phoneme correspondence) are German and Swedish, followed by French and Danish (Hedgcock & Ferris, 2009). English, on the other hand, is an extremely deep orthography (one of the deepest alphabetic orthographies), meaning that the grapheme–phoneme correspondences are much less reliable and less consistent. Decoding words is much more difficult in deep orthographies such as English, and deep orthographies require readers to rely less on letter-by-letter reading and instead to use groups of letters, morphemes, and lexical information that is unique to each word.

The deep orthography of English presents a substantial challenge to many L2 learners because of its high degree of irregularity. Many English letters can correspond to more than one sound; for example, the letter *c* can correspond to the sound [k] as in *cat* and also the sound [s] as in *certain*. On the other hand, many sounds can be represented by more than one letter in English. For example, the sound [k] can be represented by *c*, *k*, or *q*. In addition, English has a number of consonant digraphs, such as *th*, *sh*, *ch*, and *ck*, in which two graphemes are used to represent a single sound. These one-to-many and many-to-one relationships between graphemes and sounds in the orthography of English make decoding words especially difficult for learners whose first language has a shallower orthography and, thus, more regular, one-to-one relationships, such as Spanish, Italian, Serbo-Croatian, or Korean.

Another aspect of orthographic depth is the degree to which alphabetic orthographies represent vowel sounds. Languages such as English, Spanish, French, and German represent vowel sounds explicitly in their orthographies. However, Hebrew and Arabic primarily represent consonant sounds in their orthographies and short vowel sounds are not usually indicated explicitly in the orthography. As a result, L2 English learners who first developed literacy in Arabic or Hebrew may have difficulty distinguishing words that differ only in vowel sounds (e.g., *bug* and *bag*, *biscuit* and *basket*), particularly in oral reading.

Also, although English is alphabetic and thus the orthography primarily represents phonemes, due to its depth, it also strongly preserves morphological information in its orthography, often leading to increased orthographic irregularity. For example, the past tense morpheme *–ed* is orthographically consistent among the words *played*, *hunted*, and *walked*, even though it is realized phonologically in three different ways ([d], [ɪd], and [t], respectively). Similarly, the plural morpheme *–s* is preserved orthographically in *dogs* and *cats*, but not phonologically ([z] and [s], respectively). English also tends to preserve orthographic representation of word stems, even when the phonological representation changes. We see this in words such as *nation* and *national*, or *resign* and *resignation*, where the orthography maintains the spelling of the stem, even though the pronunciation shifts. This inconsistency between graphemes and phonemes could make decoding these words difficult for English learners whose L1 preserves phonological information to a greater degree.

A theory that accounts for crosslinguistic differences in reading as a result of differences in orthography is the psycholinguistic grain size theory (Ziegler & Goswami, 2005). According to this theory, children initially develop a sensitivity to larger phonological units in speech, and over time they gradually refine their sensitivity to progressively smaller units. In learning to read in their first literacy, children must figure out the optimal grain size, or the amount of orthographic information needed for efficient word decoding, for their language. Thus, in orthographically shallow languages, less orthographic information is needed for decoding and the grain size required is very small. On the other hand, in orthographically deep languages (such as English), decoding words requires more orthographic information and the grain size is much larger, such as syllables, rimes, or morphemes.

The degree to which students' prior literacy experiences will affect their English reading will depend on a number of factors. First, orthographic distance, or the degree of difference in the orthography of the L1 and L2, will impact the rate at which learners' L2 decoding skills develop, with less distance resulting in greater facilitation of L2 literacy development (Koda, 2008). Second, although many L2 English learners may have literacy experience in their L1, some may not. In this case, the absence of metalinguistic insights developed in the learner's L1 will mean that even if the learner's L1 is orthographically close to English, the learner will nonetheless need to start their English literacy development by learning the alphabetic principle.

Pedagogical implications

Generally, instruction targeting grapheme-sound correspondence is not included in L2 reading instruction for advanced learners, as they are likely to have already developed efficient word decoding processes. However, such instruction is beneficial for younger learners or adults in the early stages of English literacy development, or learners who do not have previous experience with an alphabetic orthography (Grabe, 2009).

Many teachers of reading understand that learners whose first literacy language uses a different orthography from English will face challenges in learning the orthography of English. However, many teachers underestimate these challenges, and perceptions of these challenges are sometimes clouded by other aspects of the languages, such as scripts. Learning a new orthography is not just a matter of learning a new set of symbols. Rather, it is learning a new way of understanding visual information and how it corresponds to phonological information. For L1 Chinese learners of English, a major task is understanding the alphabetic principle; that is, that individual letters correspond to sounds, rather than meaning. L1 Chinese learners might initially use a more visual strategy for learning new words, possibly memorizing new words as a whole, based on their overall shape or the initial and final letters, without processing intra-word components. Such a strategy could be successful at first, when the student only needs to learn a small

number of words that are visually more distinct. However, this strategy would not be sustainable as the learner is exposed to a greater number of words that are more visually similar. Lessons for students with a morphographic background could include phonics instruction or speeded drills in which students need to differentiate between visually similar words (e.g., *bug* and *dug*, or *quite* and *quiet*).

Unlike L1 Chinese students, students who have prior literacy experience in languages that are alphabetic but use different scripts from English (such as Arabic, Hebrew, or Korean) will face a different task. The alphabetic principle will be familiar to them, and they will be more likely to utilize phonological information when recognizing words in English, rather than rely on visual information. However, the first major task for these learners will be to learn the script of English, and their time might be best spent on practicing writing words that they already know orally (or even transliterating words in their L1 or names of people or places).

On the other hand, these learners and others whose first language orthography is shallower than English (e.g., Spanish) will face different challenges. These learners may be able to decode simple words that have regular spellings quite easily, but will, over time, face more difficultly decoding words in English due to its deep orthography. Students from orthography backgrounds that have more regular, one-to-one correspondences between graphemes and phonemes may have difficulty with the many vowel sounds in English. In many shallow orthography languages, each orthographic vowel may only have a single pronunciation, so these learners are likely to have difficulty with the highly complex vowel system in English, where a single orthographic vowel can have multiple pronunciations, some of which overlap with the pronunciations of other letters. For example, such learners may have difficulty due to the multiple pronunciations of words such as *dove*, *produce*, *present*, and *tear*. Similarly, such learners may have difficulty with vowel digraphs such as in *read, boat,* or *brain*, or consonant clusters and would need more explicit instruction and practice focusing on these. Lessons for these students could include fill-in-the-blank type exercises that target specific letters (e.g., those with multiple pronunciations) or letter combinations or clusters (e.g., *This morning, I brushed my tee__.* or *We cooked the meal in the ki___en*).

Phonics instruction is useful for helping learners become more familiar with the orthography of English. Phonics instruction could focus on individual grapheme–phoneme correspondences, or it could focus on letter clusters, as described above. Instruction could then include reading texts that make use of the sounds that the learners had learned. Some useful resources for classroom activities targeting phonics can be found in Blevins (1997, 1999). The *Words Their Way* series (Helman, Bear, Invernizzi, Templeton, & Johnston, 2012, 2014) also provides a number of practical resources that may be useful for teachers of English learners of various L1 backgrounds.

As with all reading skills, it is important for learners to develop automaticity. To develop automatic processing of orthography, learners need a high amount of exposure to print and many opportunities to read words both in isolation and in context.

In addition to the above suggestions, teachers can also become familiar with their students' literacy backgrounds. A teacher could find out what other language(s) their students speak, and the extent to which the students have literacy experience in those languages. Of course, a teacher may not be able to develop in-depth knowledge about every first language that is spoken in their classroom, but knowing whether or not students have literacy in those languages, and, if they do, the writing systems and orthographic depth of those languages can be useful information for understanding the difficulties that a student might encounter. For teachers who are interested, a recent volume, Verhoeven and Perfetti (2017), provides in-depth descriptions of the scripts and orthographies of many languages from around the world, and Joshi and Aaron (2016) provides descriptions of literacy development in many different languages, including English.

SEE ALSO: Beginning-Level Readers; First Language and Second Language Reading; Initial Literacy Development for Learners of English; Lower Primary School Readers (K-3): Foundational Knowledge and Skills for Second Language Reading; Role of Oral Language in the Development of L2 Literacy Skills

References

Akamatsu, N. (2003). The effects of first language orthographic features on second language reading in text. *Language Learning*, *53*, 207–31.

Blevins, W. (1997). *Phonemic awareness activities for early reading success*. New York, NY: Scholastic.

Blevins, W. (1999). *Quick and easy learning games: Phonics*. New York, NY: Scholastic.

Coulmas, F. (2003). *Writing systems*. New York, NY: Cambridge University Press.

Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. New York, NY: Cambridge University Press.

Hedgcock, J. S., & Ferris, D. R. (2009). *Teaching readers of English*. New York, NY: Routledge.

Helman, L., Bear, D. R., Invernizzi, M., Templeton, S., & Johnston, F. (2012). *Words their way with English learners: Word study for phonics, vocabulary, and spelling*. New York, NY: Pearson.

Helman, L., Bear, D. R., Invernizzi, M., Templeton, S., & Johnston, F. (2014). *Words their way: Within word pattern sorts for Spanish-speaking English learners*. New York, NY: Pearson.

Joshi, R. M., & Aaron, P. G. (Eds.). (2016). *Handbook of orthography and literacy*. New York, NY: Routledge.

Katz, L., & Frost, R. (1992). The reading process is different for different orthographies: The orthographic depth hypothesis. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 67–84). Amsterdam, Netherlands: Elsevier.

Koda, K. (2008). Impacts of prior literacy experience on second language learning to read. In K. Koda & A. M. Zehler (Eds.), *Learning to read across languages: Cross-linguistic relationships in first and second language literacy development* (pp. 68–96). New York, NY: Routledge.

Perfetti, C. A. (2003). The universal grammar of reading. *Scientific Studies of Reading*, *7*, 3–24.

Verhoeven, L., & Perfetti, C. (Eds.). (2017). *Learning to read across languages and writing systems*. New York, NY: Cambridge University Press.

Wang, M., Koda, K., & Perfetti, C. A. (2003). Alphabetic and non-alphabetic L1 effects in English semantic processing: A comparison of Korean and Chinese English L2 learners. *Cognition*, *87*, 129–49.

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin*, *131*, 3–29.