On the Anglocentricities of Current Reading Research and Practice: The Perils of Overreliance on an “Outlier” Orthography

David L. Share
University of Haifa

In this critique of current reading research and practice, the author contends that the extreme ambiguity of English spelling–sound correspondence has confined reading science to an insular, Anglocentric research agenda addressing theoretical and applied issues with limited relevance for a universal science of reading. The unique problems posed by this “outlier” orthography, the author argues, have focused disproportionate attention on oral reading accuracy at the expense of silent reading, meaning access, and fluency, and have significantly distorted theorizing with regard to many issues—including phonological awareness, early reading instruction, the architecture of stage models of reading development, the definition and remediation of reading disability, and the role of lexical–semantic and supraparalexical information in word recognition. The dominant theoretical paradigm in contemporary (word) reading research—the Coltheart/Baron dual-route model (see, e.g., J. Baron, 1977; M. Coltheart, 1978) and, in large measure, its connectionist rivals—arose largely in response to English spelling–sound obtuseness. The model accounts for a range of English-language findings, but it is ill-equipped to serve the interests of a universal science of reading chiefly because it overlooks a fundamental unfamiliar-to-familiar novice-to-expert dualism applicable to all words and readers in all orthographies.

Keywords: reading, word recognition, orthography, Anglocentrism

Many of the ideas elaborated in this article were first presented at a symposium dealing with universal and language-specific aspects of developmental dyslexia chaired by Brian Byrne that took place at the World Congress of Psychology in Stockholm, Sweden, in August 2000. These ideas were further developed and put to paper while I was a visiting researcher at the Faculty of Education at the University of Amsterdam in 2005. I wish to acknowledge the generous hospitality of Aryan van der Leij and Peter de Jong, and the many valuable discussions that took place during this visit. I am also indebted to Max Coltheart, Peter de Jong, Aryan van der Leij, and Tami Katzir for valuable comments on an earlier version of this article. Special thanks to Liat Bubul for assistance in preparing the article. A number of the Hebrew-language studies reported in the article were funded by the Israel Science Foundation and the Ministry of Education: Office of the Chief Scientist.

Correspondence concerning this article should be addressed to David L. Share, Department of Learning Disabilities, Faculty of Education, University of Haifa, Mount Carmel, 31905, Haifa, Israel. E-mail: dshare@construct.haifa.ac.il

1 It should be stressed that I am not espousing the view (see, e.g., Finnegan, 1987; F. Smith, 1978) that English spelling is chaotic. On the contrary, I have argued elsewhere that English is sufficiently regular to be a functional (i.e., decipherable) alphabet even for novice readers (see Share, 1995, pp. 160–163). It is now well established that consonantal letters in English are highly predictable, and even vowel letters display a large degree of conditional consistency once position, graphemic context, and morphemic regularities are taken into account (Carney, 1994; Kessler & Treiman, 2001; Venezky, 1970). Only a small proportion of the English lexicon represent isolated exceptions, such as broad and choir (Crystal, 2003), but perhaps because many of these are extremely common words (e.g., was, of, some), the reader easily gains the impression that much of printed English is arbitrary or chaotic. Nevertheless, both psychologists and linguists (see, e.g., Borgwalder et al., 2005; Daniels & Bright, 1996; Seymour et al., 2003; Ziegler et al., 1996) are unanimous in proclaiming that the English spelling–sound code (or rather cipher) is the most complex of all the world’s alphabetic orthographies.
field; rather, the field as a whole, Anglophone and non-Anglophone, is unwittingly a victim of the peculiarities of English orthography.

The sheer volume and near-unrivalled preeminence of English-language reading research and theorizing, together with the proliferation of the Anglophone research agenda around the globe, have led a growing number of theorists to question the applicability of Anglophone findings to other languages and orthographies (e.g., Perfetti, Liu, & Tan, 2005; Seymour, 2005; Ziegler & Goswami, 2005). Recent evidence from fields such as dyslexia (e.g., Goulandris, 2003; Vellutino, Fletcher, Snowling, & Scanlon, 2004), reading development (e.g., Aaron & Joshi, 2006; Harris & Hatano, 1999; Ziegler & Goswami, 2005), and skilled reading (Frost, 1998, 2005; Perfetti, Liu, & Tan, 2005) has led to expressions of unease regarding the exceptional nature of English orthography in comparison with other alphabetic orthographies. In the developmental literature, this deviance has been characterized as “dramatic” (Frith, Wimmer, & Landerl, 1998, p. 39; Hutzler, Ziegler, Perry, Wimmer, & Zorzi, 2004, p. 273), “extreme” (Seymour, 2005, p. 310), “profound” (N. C. Ellis & Hooper, 2001, p. 589), and “particularly unnatural” (Snowling & Hulme, 2005, p. xiv). These qualms, however, have been voiced in most cases as a footnote to the voluminous English-language research literature.

In the current critique, I place the issue of generalizability at the center of a detailed consideration of the repercussions of English orthographic exceptionality for the scope and prioritization of today’s reading research agenda, in general, and developmental word reading research, in particular. I begin this critique with the question: Is English really that exceptional?

**English: An “Outlier” Orthography**

In the most ambitious cross-linguistic study of beginning reading launched to date, Seymour, Aro, and Erskine (2003) found that most children from the 14 participating nations were reasonably accurate and fluent decoders by the end of the 1st school year, averaging 87% accuracy for familiar high-frequency words. The English word-reading result, 34%, was over three standard deviations below the 14-nation mean (the 87% average included the English figure) and far below the next poorest outcome—Danish (71%). These dramatic differences do not appear attributable to factors such as age, gender, syllable complexity, or letter knowledge. Perhaps even more astonishing were differences in the nonword results, 29% accuracy for English compared with the cross-national average of 82%. It is a telling observation that in many analyses, the English data were treated as an outlier or either excluded or analyzed separately. An important issue in evaluating findings, such as those described above, is the commensurability of data based on samples drawn from different cultures, from school systems with divergent curricula, and from different sets of words.

**Item Commensurability**

One obvious challenge for any cross-linguistic study is to match test stimuli on factors such as length, frequency, syllable structure, and so forth. When test items are individually matched across linguistic cousins, such as English and German, using cognates (three/drei, garden/Garten) or even identical words (Land, Ball), the same dramatic differences are seen as in Seymour et al.’s (2003) study (Frith et al., 1998; Landerl, Wimmer, & Frith, 1997; Wimmer & Goswami, 1994). Even pseudowords that call only for application of simple grapheme–phoneme correspondences reveal the same decrements among English readers (Frith et al., 1998; Landerl et al., 1997; Wimmer & Goswami, 1994). Frith et al. (1998), for example, compared identical nonwords (e.g., twivcan) in the two languages; at age 8 years, the incidence of errors among English speakers was 30%, compared with only 0.5% for German speakers. Each of the above methods for maximizing item comparability has its limitations (see discussion in N. C. Ellis et al., 2004), yet the same pattern of findings emerges whichever approach is adopted. These data, therefore, do not appear to be an artifact of methodology.

A highly innovative approach to matching items across languages was developed by N. C. Ellis and Hooper (2001). They compared Welsh and English using frequency-matched lists compiled by sampling words from 100 successive strata of decreasing written word frequency in each language, each successive item representing a successively lower frequency 10,000 word band. All other variables—such as length, syllable structure, meaningfulness, and so forth—were left free to vary. Oral reading accuracy among English speakers/readers reached the 52% mark, compared with 61% for Welsh. This indicates that the Welsh group (ages 6 and 7 years) were able to read twice as many words in their printed language than were English readers. In their Welsh–English study, Spencer and Hanley (2003) reported a similar outcome using conventional item-matching methods (cognates and derived nonwords); hence, it appears that the extreme deviance of English is robust across methodologies.

**Subject Commensurability**

N. C. Ellis and Hooper’s (2001) study, which is mentioned above, matched their Welsh-medium and English-medium samples by catchment area, by classroom size, and by reading curriculum and instructional time. Even after controlling for these important instructional factors, the magnitude of the between-languages differences again replicated Seymour et al.’s (2003) data, with the English group performing only half as well as the children learning the more regular Welsh orthography. Similar results were reported in Spencer and Hanley’s (2003) Welsh–English study, in Landerl’s (2000) German–English investigation, and Oney and Goldman’s (1984) Turkish–English comparison, all of which matched groups on reading program.

The most convincing evidence that sociocultural differences are not responsible for the poor English outcomes comes from within-subject bilingual comparisons. Geva and colleagues (Geva & Siegel, 2000; Geva, Wade-Woolley, & Shany, 1993) compared English and Hebrew reading performance in a large sample of native English-speaking Canadian children who were learning highly regular pointed Hebrew for 2–3 periods a day at an English–Hebrew bilingual school. These children’s oral word reading ac-

---

2 The terms pseudoword and nonword are used interchangeably throughout this article to refer to invented (i.e., possible) word-like strings of letters.
curacy in Grade 1 far outstripped their own native English ability despite limited oral L2 Hebrew proficiency. Even in Grade 5, word recognition accuracy in L1 English was inferior to L2 Hebrew in Grade 1.

Finally, additional support for the claim that orthographic inconsistency is the source of difficulty in English reading acquisition, rather than item or subject confounds, comes from comparisons between monolingual English speakers taught with a regularized pedagogy, such as the initial teaching alphabet (i.t.a.), or with conventional spelling (Downing, 1967; Thorstad, 1991). Thorstad (1991), for example, found that performance decoding i.t.a. was superior to conventional (English) orthography and similar to Italian readers matched for age and nonverbal ability.

Although no individual study has controlled for all relevant item and subject variables, the body of evidence, collectively, is unanimous in showing that, for the developing reader, English is truly exceptional. By the end of the 1st year of schooling, hyperlexic-style reading is the norm in transparent alphabetic orthographies: most children are capable of tackling almost any printed (monosyllabic) word. In English, such proficiency is delayed for several years. Moreover, this “great divide” between early English and non-English reading appears to be more than quantitative and clearly extends to the nature of the reading strategies employed.

English Orthography Is Abnormal Qualitatively and Quantitatively

Among young readers, a variety of qualitative differences in reading strategies has been discovered. Compared with English, length and regularity effects are significantly greater in shallower orthographies with a higher incidence of neologisms—mispronunciations resulting in nonwords (e.g., N. C. Ellis & Hooper, 2001; Seymour et al., 2003; Spencer & Hanley, 2003). English readers, in contrast, tend to make more real-word substitutions (termed lexicalizations; e.g., shoe for school; N. C. Ellis & Hooper, 2001; Frith et al., 1998; Geva & Siegel, 2000; Landerl et al., 1997; Seymour et al., 2003; Spencer & Hanley, 2003; Thorstad, 1991; Wimmer & Goswami, 1994), exhibit frequent nonresponses or refusals (N. C. Ellis & Hooper, 2001; Seymour et al., 2003; Spencer & Hanley, 2003), reduced word–nonword (lexicality) effects (Landerl et al., 1997; Patel, Snowling, & de Jong, 2004; Seymour et al., 2003; Wimmer & Goswami, 1994), and stronger pseudohomophone effects (Goswami, Ziegler, Dalton, & Schneider, 2001). This latter finding indicates greater reliance on whole word phonology in nonword reading.

Observations of the decoding process also underscore qualitative differences. Readers of shallow orthographies compared with readers of English display more exhaustive (and often laborious) letter-by-letter decoding (Landerl, 2000; Thaler, Ebner, Wimmer, & Landerl, 2004; Thorstad, 1991), and this finding is supported by eye-movement studies (e.g., Hutzler & Wimmer, 2004; Zoccolotti et al., 1999) and brain imaging studies (e.g., Paulesu et al., 2000).

The reliability and persistence of these qualitative differences even among skilled adult readers (a thorough review is beyond the scope of the present article) have inspired several prominent theories concerned with the (different) reading processes invoked in deep versus shallow orthographies (see Ziegler & Goswami’s, 2005, psycholinguistic grain-size theory, and the orthographic depth hypothesis proposed by Katz & Frost, 1992, and Frost, 2005). Each of these theories bears witness to the fact that the early quantitative and qualitative differences discussed above are not transitory phenomena but leave developmental “footprints” (see Ziegler & Goswami, 2005) across the entire reading lifespan.

Why Is English Unique Among Writing Systems?

Grammatologists agree that English underwent a series of unparalleled historical changes. First, Christian missionaries borrowed an orthography designed to represent far fewer than its dozen or so vowels. English then absorbed successive waves of invaders, conquerors, and borrowings before fossilizing spellings at the advent of printing and on the threshold of the Great Vowel Shift, yet repudiating almost all attempts at spelling reform. The upshot of these successive upheavals is an amalgam of subsystems of spelling (principally Germanic, Norman-French, and Latin-Greek) that provide a splendid “fossil record” of the geopolitical and cultural history of the English language (Carney, 1994; Crystal, 2003; Scragg, 1974). Unfortunately for the novice reader, however, it strays unusually far from the one-letter-one-phoneme mapping principle that is the norm among the world’s writing systems (Daniels & Bright, 1996). In particular, the major source of irregularity is the set of only five vowel letters (six if we include y) representing about 20 vowel phonemes (letters representing consonants are much less ambiguous). This mismatch between the number of letters and the number of phonemes is reflected in the fact that most proposals for spelling reform have called for increasing the number of letters or adding diacritics (see Carney, 1994, chapter 7; Scragg, 1974, chapter 6).

The history of English spelling reveals the source of this “bewildering variety” of spelling conventions (see Scragg, 1974, p. 60) but does not indicate the degree to which English differs from other orthographies. Similar historical events have occurred in the history of other orthographies; it is doubtful that any one of these cultural/historical factors is unique to English (see, e.g., Elbro, 2006). The uniqueness of English may arise from the combination and/or magnitude of these factors.

Reading researchers have yet to agree on a principled way of determining precisely how English differs from other writing systems. Attempts to characterize the complexity of English spelling–sound relations reveal little convergence, differing with regard to the appropriate orthographic unit of analysis (letter, grapheme, onset/vowel/coda, or rime [vowel-plus-coda]), corpora (e.g., monosyllabic vs. polysyllabic), quantification methods, and so on. Even the terms regular/irregular and consistent/inconsistent commonly employed in the English-language reading literature embody different processing assumptions regarding the psychological reality of the mapping relation. Only three studies (Borgwaldt, Hellwig, & de Groot, 2005; van den Bosch, Content, Dalemans, & de Gelder, 1994; Ziegler, Jacobs, & Stone, 1996) have directly compared orthographies on a common metric.

Ziegler et al. (1996) tallied the number of monosyllabic words (types) in which the spelling “body” (vowel nucleus and coda)

---

3 For a discussion of how the beginning reader copes with the challenges posed by this complexity, see the discussion of the notion of lexicalization in Share (1995, pp. 163–165; see also Elbro, 2006) as well as Glaser’s (1984) concept of pedagogical theories.
either does or does not map onto one and only one phonological body (cf. -ack vs. -aid). English was found to be more “feedforward” (spelling-to-sound) inconsistent than French (31% vs. 12%). Borgwaldt et al. (2005) compared word-initial letter-to-phoneme consistency in several large corpora of monosyllabic and polysyllabic lemmas expressed as (continuous) entropy values in seven European languages. Overall entropy values were highest in English, with an average of 3.89 phonemes per word-initial letter. The next highest value (2.85) was French. Consonantal entropy values in English were slightly lower than French and comparable with German, whereas the value for vowels (as previously noted) was almost twice that of the next most ambiguous script, German.

Van den Bosch et al. (1994) developed a two-dimensional measure of orthographic complexity representing the intersection of (1) the complexity of letter–phoneme alignment (graphemic parsing) and (2) the complexity of grapheme–phoneme correspondence. Several automatic, language-independent, data-driven algorithms were trained on a subset of words from 20,000-word corpora (monosyllabic and polysyllabic) in each of three languages: English, French, and Dutch. The algorithms were then tested for generalization to the remainder of the corpus. In terms of grapheme-to-phoneme mappings, English was shown to be the most complex (i.e., irregular), by far, but somewhat less complex than French and similar to Dutch in the complexity of graphemic parsing. Given that neither linguists nor psychologists are able to agree on the method of analysis (even within English), it is hardly surprising that we lack a consensual cross-linguistic metric. Part of the problem is that studies of spelling–sound consistency often derive from very different research agendas.

To summarize, English presents a fascinating case study of orthographic liberalism-turned-conservatism. Unfortunately, for the cause of reading science it also represents the most extreme case of spelling–sound complexity; on this latter point, there is no dispute.

Dual-Route Theory and the Challenges of Irregularity

The distinction between words that conform to the rules of English grapheme-to-phoneme correspondence (regular words and nonwords) versus those that do not conform (irregular or exception words) has been the main inspiration for the influential dual-route model of skilled word reading (M. Coltheart, 1978, 2001; M. Coltheart, Curtis, Atkins, & Haller, 1993; M. Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001) and its rivals (Harm & Seidenberg, 1999; Perry, Ziegler, & Zorzi, 2007; Plaut, McClelland, Seidenberg, & Patterson, 1996; Seidenberg & McClelland, 1989; Woollams, Ralph, Plaut, & Patterson, 2007; Zorzi, Houghton, & Butterworth, 1998). The dual-route model has been acclaimed “The most influential theory of visual word recognition” (G. Lukatela & Turvey, 1998, pp. 1059–1060), and it ranks as one of the most prolific theories in modern cognitive psychology. The dual-route model has not gone unchallenged (see, e.g., Glushko, 1979; Humphreys & Evett, 1985; Seidenberg & McClelland, 1989) but still enjoys virtual benchmark status in theorizing about skilled word recognition (M. Coltheart, 2005; M. Coltheart et al., 1993, 2001; Perry et al., 2007; Zorzi et al., 1998), reading development (e.g., Byrne, Freebody, & Gates, 1992; Jackson & Coltheart, 2001; Jorm & Share, 1983; Seymour, 2005), developmental and acquired dyslexia (e.g., Castles & Coltheart, 1993; Manis, Seidenberg, Doi, McBride-Chang, & Petersen, 1996; Stanovich, Siegel, & Gottardo, 1997), and even spelling (e.g., Barry, 1994; A. W. Ellis & Young, 1988). Neuro-imaging work in the field of word recognition has also been dominated by the dual-route approach (e.g., Jobard, Crivello, & Tzourio-Mazoyer, 2003; Price et al., 2003).

The central axiom of the Coltheart/Baron version of the dual-route model (see, e.g., Baron, 1977; M. Coltheart, 1978) is that no single procedure yields correct pronunciations of both nonwords (e.g., slint) and exception words (e.g., pint). Nonwords can only be correctly pronounced via grapheme–phoneme correspondence rules (the “nonlexical” route); exception words require an additional procedure (the “lexical” route) because they cannot be pronounced by the rules.

What we mean by the term dual-route model is a model that has a route that can read words but cannot read nonwords and another route that can read nonwords and regular words but misreads exception words by regularizing them. (M. Coltheart et al., 1993, p. 590)

The dual-route model unequivocally rejects the connectionist counterclaim (Plaut et al., 1996; Seidenberg & McClelland, 1989) that the human reading system contains a single processing procedure that correctly pronounces both nonwords and exception words.

Regular words can be read by either of the two routes; thus, the nonword/exception-word contrast is decisive within the (Coltheart/Baron) dual-route framework. The critical nonword/exception contrast lies at the heart of all but one of six questions that, according to M. Coltheart et al. (1993), “any serious model might be expected to supply answers to” (p. 590). These included the following: “How do skilled readers read exception words? (Question 1); how do skilled readers read nonwords? (Question 2); and how do acquired and developmental exception/nonword dissociations arise? (Questions 4, 5, and 6). The nonword/exception contrast was also the foremost concern in setting parameters in the 2001 implementation of the dual-route model.

The most delicate issue with the DRC [dual-route cascaded] model is to try and set an appropriate balance between the two routes of the model . . . What is needed is a set of parameters under which DRC reads all exception words and all nonwords correctly. (M. Coltheart et al., 2001, p. 219)

M. Coltheart et al. (2001) evaluated the model by comparing its performance to standard effects obtained by skilled readers only after parameter-setting enabled the model to read both members of exception/nonword pairs, such as chef/starm. The centrality of the nonword/exception contrast is also unmistakable in the principal counterclaim of the rival connectionist approach:

A key feature of the model we propose is the assumption that there is a single, uniform procedure for computing a phonological representation from an orthographic representation that is applicable to exception words and nonwords as well as regular words. (Seidenberg & McClelland, 1989, p. 525)

The connectionist (or “triangle”) model developed by Seidenberg and colleagues (Plaut et al., 1996; Seidenberg & McClelland, 1989) evolved largely as an alternative to the traditional Coltheart/
Baron dual-route framework; thus, the same empirical legacy that has dominated dual-route investigations (namely, the nonword/exception contrast) has also pervaded much of connectionist decision-making regarding architecture, implementation, and evaluation. Indeed, the crucial failing of Seidenberg and McClelland’s (1989) first-generation model, and the principal motive for the model’s first major overhaul, was its inability to successfully pronounce nonwords while maintaining the ability to pronounce both exception and regular words (see Besner, Twilley, McCann, & Seergobin, 1990; Plaut et al., 1996). Only recently has the Seidenberg group moved on to address issues (e.g., meaning access; see Harm & Seidenberg, 2004) other than the oral pronunciation of isolated words differing in spelling–sound consistency.

When Irregularity Is the Exception

If a writing system contains no exception words, is a second route necessary? Most orthographies are fairly regular in terms of print-to-sound relations (Daniels & Bright, 1996; Seymour et al., 2003); thus, a single rule-based mechanism should be adequate for pronouncing all (or nearly all) letter strings. A growing number of reading researchers have begun to question the generalizability of the dual-route architecture beyond English (e.g., Bishop & Snowling, 2004; Hutzler & Wimmer, 2004; Ziegler & Goswami, 2005). Ziegler and Goswami (2005), for example, ponder whether “it might even be the case that the prominent dual route architecture (i.e., two separate routes in the skilled reading system) may in fact only develop for English” (p. 15). Even the dual-route theorists themselves appear to share these misgivings in attempting to extend the DRC model to German, a highly regular script:

[M.] Coltheart et al.’s (1993) major justification for a dual-route architecture in English was the fact that neither a lexical nor a nonlexical system alone seemed sufficient to explain both nonword and exception word reading, as well as the double dissociation between surface and phonological dyslexia. However, if a major part of German words (about 95% if loan words are excluded) can be pronounced correctly using a limited set of general rules, one might argue that there is little need for a separate lexical system. (Ziegler, Perry, & Coltheart, 2000, pp. 426–427)

If spelling–sound regularity is the global norm, then a second route tailored specifically to irregular words is dispensable. A more parsimonious one-route model would deal satisfactorily with words and nonwords alike.

Ironically, these reservations about the lexical route, much like the Anglocentric regular/irregular dichotomy in the dual-route model itself, overlook a somewhat different, yet fundamental, duality in word reading that applies to all words in all possible orthographies regardless of their degree of “regularity.” On the one hand, all words are visually unfamiliar at some point in reading development; thus, the reader must possess some algorithm, albeit imperfect, for independently identifying words encountered for the first time (in context; see Share, 1995, for a more detailed discussion).4 On the other hand, the reader must eventually achieve a high degree of automatization in word recognition—rapid and effortless recognition of familiar words and morphemes (LaBerge & Samuels, 1974; Logan, 1988, 1997; Perfetti, 1985) perceived as whole units via a direct-retrieval mechanism (see Ans, Carbonnel, & Valdois, 1998; Weekes, 1997). This ability to automatize or “modularize” word identification (Adams, 1990; Stanovich, 1990, 2000), far more than the ability to derive the correct pronunciations of words varying in spelling–sound consistency, is probably the quintessence of reading skill (Perfetti, 1985). As with the decoding algorithm, this high-speed, direct-retrieval mode applies to all words in all orthographies, regular and exceptional.

This universal dualism merges the study of reading with the study of human skill learning across a range of domains (see, e.g., J. R. Anderson, 1981; Karni, 1996; LaBerge & Brown, 1989; Logan, 1988; Newell & Rosenbloom, 1981; Shiffrin & Schneider, 1977; Siegler, 1988). The dualism common to almost all skill learning is a contrast between, or transition from, slow, deliberate, step-by-step, unskilled performance to rapid, automatized, one-step or “holistic” skilled performance. Without this transition, the “skill” of reading might never have influenced modern knowledge-based cultures so profoundly (D. R. Olson, 1994).

This basic and universal “novice-to-expert” dualism (from a reader’s perspective) or “unfamiliar-to-familiar” dualism (from an item-based perspective) also converges with the dual nature of an efficient orthography. Specifically, an efficient script can be conceptualized as a compromise between the competing needs of the novice and the expert reader (Rogers, 1995; Venezky, 2007). This orthographic dualism might be termed the “decipherability/automatizability” criterion.

An effective orthography must provide the reader with a means for deciphering new words independently (see Share, 1995). This applies to the young child new to the world of print and to the skilled reader encountering a new or unfamiliar word. Furthermore, and this is crucial to skill learning in all domains, this algorithmic process must lay foundations for rapid, direct-retrieval mechanisms. This “do-it-yourself” or “self-teaching” function of decoding (Jorm & Share, 1983; Share, 1995, 2008) is probably the chief virtue of alphabetic scripts, supplying an economical means for identifying new words (via print to sound translation) and establishing the detailed orthographic representations on which rapid, fully-unitized skilled word recognition is founded.

A successful script must also answer the needs of experts by providing distinctive word-specific (or morpheme-specific) visual–orthographic configurations required for unitizing and automatizing skilled word recognition. Ideally, each morpheme should have one and only one representation (morpheme “constancy”) without showing morphophonemic variation (e.g., electric/electricity/electrical), with different morphemes represented purely by rote whole-word methods (for discussion, see Share, 1995)—there are simply too many words. This is probably the main reason no pure logography has ever existed or, indeed, could ever exist (Gelb, 1952; Mattingly, 1985). Even in Chinese, characters that directly represent meanings constitute only a tiny fraction of the Chinese character corpus (Taylor & Taylor, 1995). The overwhelming majority of Chinese words are represented by compounds containing elements indicating meaning (the radical) and sound (phonetic). Even when taught without the support of supplementary alphabetic scripts, such as pinyin, the combination of semantic and phonetic information appears to be highly productive for the novice (see R. C. Anderson, Li, Ku, Shu, & Wu, 2003; Cheung et al., 2006; Shu, Anderson, & Wu, 2000).
differently (morpheme “distinctiveness”; Rogers, 1995). Historically, the morphemic principle, essential for automatizability, has taken precedence over the competing decipherability principle (Daniels & Bright, 1996; Perfetti, 2003; Venekly, 2007; but see Serbo-Croatian; K. Lukatela, Carello, Shankweiler, & Liberman, 1995). Word separation may also be a crucial factor in the unitization/automatization process (Saenger, 1991).

A script that caters primarily to the needs of skilled readers, such as precomunist Chinese characters (and in many respects English orthography; see Chomsky & Halle, 1968), will pose enormous challenges for novices. Conversely, a script that provides maximum decipherability for novices—the highly regular pedagogies, such as i.t.a., Korean hangul, or Japanese kana—will often fail (as a stand-alone script) to meet the needs of skilled readers, primarily owing to homophony.

Recognizing the divergent decipherability/automatizability needs of learners and experts, some scripts have dual versions. Among these are the pointed (fully voweled) and unpointed (incompletely voweled) variants of Hebrew (Ravid, 2006), Arabic (Abu-Rabia & Taha, 2006), and Farsi (Baluch & Besner, 1991), and the fully briscriptual systems in China (pinyin/characters), Korea (hangu/hancha), and Japan (Kanjikana). In these languages, separate phonemic (or syllabic/moraic) and morphemic orthographies fulfill the two decipherability/automatizability functions.

**Where Are the Jabberwockies?**

It is important to emphasize that, unlike the standard contrast between different types of items (regular/irregular, exception/nonword, or word/nonword) on which the Coltheart/Baron dual-route conception was founded, the unfamiliar-to-familiar dualism advanced here is not a between-items contrast differentiating familiar words on the one hand and unfamiliar on the other (or high-frequency vs. low-frequency words). It is, first and foremost, a within-item developmental transition from unfamiliar to familiar (hence the phrasing “unfamiliar-to-familiar” rather than “unfamiliar-and-familiar”). This broad dualism underscores the fact that every printed word is unfamiliar at some point in development (even an individual’s own name) and calls for the application of some sort of decoding/learning algorithm. That is, every letter string is functionally a nonword at first. The Czech string *šer* (my own surname) was functionally an unfamiliar pseudoword until quite recently. Many real but rare words, such as *thegn* and *veldt* (along with some 4,000 other low-frequency items excluded by Kessler & Treiman, 2001, in their corpus of monosyllabic words deemed to be familiar to skilled adult readers), are functionally nonwords for the vast majority of readers. Conversely, any pseudoword, like any real word that has yet to be encountered in print, may become a familiar, well-unitized letter string. A pseudoword, such as *Jabberwocky*, will be familiar to many literate English speakers/readers (pseudowords are a common everyday occurrence in personal and place names, [e.g., Clinton, London], commercial brand names [e.g., Colgate, Toyota], children’s verse, etc.). Indeed, for the author, *Jabberwocky* has been familiar for much longer than *šer* or *thegn*. Thus, this unfamiliar/familiar dualism applies to all words in all possible orthographies (neither regularity/consistency nor lexicality is an issue here). A reading architecture in which distinct routes or routines are dedicated to specific types of items misses this overarching dualism entirely.

**A Dual-Route Duel . . . and Rout**

In keeping with this broad, universal dualism, the earliest formulations of the dual-route model focused on the unfamiliar/familiar distinction, ignoring the regularity dimension. In his brief chronicle of dual-route history, M. Coltheart has credited Saussure (1922) with the first articulation of a dual-route conception of reading: “We read in two ways; the new or unknown word is scanned letter after letter, but a common word is taken in at a glance” (M. Coltheart, 2005, p. 6). The focus here is clearly on the familiarity dimension.

One of the first modern cognitivist formulations of the dual-route notion (Forster & Chambers, 1973) adopted this same approach. Forster and Chambers (1973) discussed two alternative ways to pronounce printed words:

First, the pronunciation could be computed by application of a set of grapheme–phoneme rules, or letter-sound correspondence rules. This coding can be carried out independently of any consideration of the meaning or familiarity of the letter sequence . . . Alternatively, the pronunciation may be determined by searching long-term memory for stored information about how to pronounce familiar letter sequences, obtaining the necessary information by a direct dictionary look-up, instead of rule application. Obviously, this procedure would work only for familiar words. (p. 627)

Once again, the key dualism here, is familiarity, not regularity. Accordingly, Forster and Chambers (1973) focused their experimental manipulations on familiarity (operationalized via printed word frequency), comparing words with nonwords and high-frequency with low-frequency words. Both high-frequency (e.g., *knee*, winter) and low-frequency (e.g., *dank*, ostler) items included.

---

5 There are clearly advantages for a script that maintains morpheme “constancy”—the same morpheme always written the same way—but it may be morpheme distinctiveness that is crucial for the automatization of word recognition. In other words, it is not that the *w* in *two* is important for revealing morphemic relationships (e.g., *twelve*, twice, twilight)—a highly doubtful assumption for the young reader—but that this etymological quirk provides unique spellings for potentially confused homophones (*toolw*/*to*/). (Historically, conscious efforts were often made by spelling reformers to avoid homophones becoming homographs and thereby remove possible ambiguity (Carney, 1994, chapter 7; Scragg, 1974, chapter 4).

6 Purely phonemic scripts (such as pinyin and Zhi-Yin-Fu-Hao) as well as syllabic/moraic scripts (such as Japanese kana) appear to be remarkably easy to learn to decode (e.g., Mason, Anderson, Onnara, Uchida, & Imai, 1989; McCarthy, 1995; Taylor & Taylor, 1995), but if extensive homophony exists in the spoken language, such scripts will violate the morphemic distinctiveness principle and prove impracticable without supplementary morpheme-based characters (as in the case of Japanese Kanji and Korean Hancha) that provide more direct links to morpheme identity. By the same token, *š*/*š* was never destined to supplant conventional English orthography because it ignores the morphemic principle. Conversely, primarily morphemic scripts—such as Chinese characters that are well-adapted to serving the literacy needs of a small highly educated elite but require the novice to invest extraordinary amounts of time and effort (Hoosain, 1995; Taylor & Taylor, 1995)—are far more accessible to the novice (and the general populace) when supplemented by decipherable phonemic scripts, such as pinyin (Siok & Fletcher, 2001; Taylor & Taylor, 1995).
a mixture of both “regular” and “exception” words. Regularity was not an independent variable nor was it even mentioned in their article.

As M. Coltheart (2005) observed, an emphasis on known and unknown words focuses attention on familiarity. Hence, the main contrast is between words that, according to the dual-route model, can be read by a direct- retrieval (or “lexical” routine) and non-words that cannot be retrieved directly and therefore depend on a nonlexical procedure. According to M. Coltheart (2005), “Baron and Strawson (1976) were the first to see that, within the context of dual-route models, this is not quite the right contrast (at least for English) [italics added]” (p. 7). Baron and Strawson sought evidence for rule use when skilled readers pronounced words aloud. Hence, their experimental manipulations focused on words that varied in degree of rule accessibility. They compared (a) “regular” words, which follow the “rules” of English orthography, (b) “exception” words, which break these rules, and (c) “nonsense” words, which can only be pronounced by the rules (p. 387). In contrast to Forster and Chambers (1973), regularity was the key experimental variable in Baron and Strawson’s investigation, with familiarity relegated to the status of a potential confound and held constant. However, many words (in English) are regular and accessible to the rules; thus, the contrast between words and nonwords is no longer adequate within the “regularity/irregularity” dual-route orientation. It becomes necessary to distinguish between regular and irregular words, the latter items becoming indispensable and familiarity (and frequency) superfluous. Neither familiarity nor frequency was a variable in Baron and Strawson’s study.

The Coltheart/Baron dual-route conception shifted focus from the original and universal familiarity-oriented dualism to the peculiarly Anglophone regular/irregular dualism and (unwittingly) promoted a particularistic English-language research agenda. This was at the expense of a dualism better adapted to serve the broad interests of reading science in which, orthographically speaking, irregularity is the exception to the rule.

Preoccupation with the problems of spelling–sound infidelity has affected the reading research outlook both theoretically and empirically. Attention has been diverted to a variety of empirical and methodological issues with limited applicability to a universal reading science. These are discussed below beginning with the issue of accuracy versus fluency.

**Accuracy and Fluency in Early Reading**

Given the prevalence of English spelling–sound irregularity, the challenge of deriving an accurate, or at least approximate pronunciation, is understandably a pressing concern for readers who are confronted with an unfamiliar letter string. If the new item is not correctly identified, the entire word-learning process is derailed. Young readers continually encounter large numbers of unfamiliar words (Foorman, Francis, Davidson, Harm, & Grifﬁn, 2004). For readers of English, many of these items will be irregular (Foorman et al., 2004).

Developmental reading research in the English language has, until recently, been dominated by measures of accuracy rather than speed or fluency. Accuracy, however, is largely a nonissue for the majority of the world’s (alphabetic) orthographies in which performance levels approach ceiling by the end of Grade 1 (Seymour et al., 2003). When accuracy asymptotes quickly, speed and fluency become the discriminating measures of developmental and individual differences (Breznitz, 1997; Cosso, 1999; de Jong & van der Leij, 2003; Leppanen, Niemi, Aunola, & Nurmi, 2006; Lytinen, Aro, & Holopainen, 2004; Nikolopoulos, Goulardris, Hulme, & Snowling, 2006; Wimmer, 1993; Yap & van der Leij, 1993; Zoccolotti et al., 1999), as is the case for skilled readers of English for whom error rates are often uninformative. To illustrate just how irrelevant accuracy can be in regular orthographies, consider Cosso’s (1999) report that reading performance in Italian reached 97.8% in a large normative survey sample in Grade 1, increasing by Grade 3 to 99.6%. Mean reading speed, in contrast, decreased from 3.5 s per word in Grade 1 to 2.1 s by Grade 3. In German (Wimmer, 1993), Dutch (Yap & van der Leij, 1993), Norwegian (Lundberg & Hoen, 1990), Italian (Zoccolotti et al., 1999), Greek (Porpodas, 2006), Finish (Lytinen et al., 2004), Hungarian (Csepe, 2006), and Hebrew (Breznitz, 1997), even dyslexics attain high levels of reading accuracy but remain slow.

**Reading Rate, Efficiency, Automaticity, and Fluency**

The speed and automaticity of sublexical and lexical processes have special significance in reading owing to the multitask demands of text comprehension and the bottlenecks created when basic word recognition processes are slow and labored (e.g., Baddeley & Gathercole, 1992; Just & Carpenter, 1987; Kintsch, 1988; LaBerge & Samuels, 1974; Oakhill, 1993; Perfetti, 1985; Swanson & Siegel, 2001). Limitations of human attention and memory are omnipresent in models of text comprehension (e.g., Just & Carpenter, 1987; Kintsch, 1988; Perfetti, Landi, & Oakhill, 2005).

Sheer speed or rate seems to relatively straightforward construct to measure (but see Carver, 1990) but is meaningless if accuracy is ignored. In certain contexts, speed is associated with haste (Carver, 1990; Frith, 1980; Hendriks & Kolk, 1997; Leinonen et al., 2001; R. K. Olson, Kliegl, Davidson, & Foltz, 1985) or with an
impulsive, nonreflective cognitive style (Kagan, Rosman, Day, Albert, & Phillips, 1964; R. K. Olson et al., 1985; Share, Jorm, Maclean, & Matthews, 1984). The concept of “efficiency,” speed adjusted for accuracy (Breznitz, 2002, 2006; Nathan & Stanovich, 1991; Perfetti, 1985; Torgesen, 2002), however, introduces a range of measurement obstacles (discussed below) that have yet to be resolved. These problems are not insurmountable when accuracy is at ceiling levels, as is the norm in most conventional orthographies, but English’s spelling–sound intractability poses a real dilemma. The problems really begin, however, when we consider the critical automaticity element that is integral to most definitions of fluency.

The general concept of automaticity is widely employed in psychology; however, there is little agreement on how to define it (Humphreys, 1985; Kahneman & Triesman, 1984; Logan, 1985; Moors & de Houwer, 2006; Stanovich, 1990). Part of the problem is that the construct is typically operationalized in complex dual-task or priming paradigms that suffer from numerous methodological and interpretational difficulties (Bargh & Chartrand, 2000; Holender, 1986; Howe & Rabinowitz, 1989; Navon & Gopher, 1980; Stanovich, 1990). In a recent review, Moors and de Houwer (2006) noted that “there is no consensus about what automaticity means” (p. 297). Stanovich (1990) has proposed replacing the troubled notion of automaticity with the concepts of modularity and information encapsulation. Logan’s (1988, 1997) instance-based notion of automaticity and allied multitrace theories of reading (see, e.g., Ans et al., 1998) have yet to find their way into mainstream reading science.

The term fluency derives from the Latin fluere—meaning “flowing”—and clearly implies more than speed and effortlessness. A news reader aims for fluency, not speed (only commercials place a premium on both rate and fluency of delivery). Accordingly, one class of fluency definitions emphasizes smooth and unbroken oral text reading operationalized as the number of pauses, hesitations, and repetitions in oral reading (e.g., Pinnell et al., 1995; Saiegh-Haddad, 2003; Zutell & Rasinski, 1991). However, a smooth, unhesitant oral rendering of a written text also falls short of the mark because it lacks the essential element of phrasing and expression manifest in variability in rhythm, juncture, pitch, stress, and duration. Some definitions of fluency regard prosody as an essential, if not the essential ingredient in fluency (e.g., Dowhower, 1991; Schreiber, 1991; Zutell & Rasinski, 1991). Stayer and Allington (1991) have gone further with a definition reminiscent of the classical arts of elocution and rhetoric. They have emphasized intonation, juncture, stress, and, above all, emotional expression, using dramatization to assess fluency. Although some prosody-oriented definitions retain criteria involving word recognition accuracy and rate (e.g., Dowhower, 1991; National Reading Panel, 2000), others jettison word-level elements altogether (Schreiber, 1991; Stayer & Allington, 1991). Wolf and Katzir-Cohen (2001) have proposed an all-inclusive definition: “[A] level of accuracy and rate where decoding is relatively effortless; where oral reading is smooth and accurate with correct prosody; and where attention can be allocated to comprehension” (p. 219). In a candid admission, the authors arrived at the “unsettling conclusion . . . that reading fluency involves every process and subskill involved in reading” (p. 220). Yet another quandary in defining fluency is whether, or how, prosody-based definitions apply to silent reading.

Another class of definitions is explicitly multidimensional (e.g., Bear, 1991; Berninger, Abbott, Billingsley, & Nagy, 2002; Katzir et al., 2006; Wood, Flowers, & Grigorenko, 2002). Berninger et al. (2002), for example, held that fluency operates differently at the word, sentence, and text levels. Katzir et al. (2006) draw a developmental distinction between fluency at the level of letters, of words, and of connected text. Bear (1991) separates word-level fluency from phrase-level fluency, whereas Wood et al. (2002) postulate an anticipatory fluency defined as preprocessing of upcoming stimuli in addition to fast and automatic recognition of letters, words, and phrases. Such componential approaches imply that readers may be fluent in some aspects of reading but not in others.

Sorely lacking is a comprehensive, empirical inquiry aimed at determining whether the multitude of features associated with the term “fluency” tap a common construct or separable abilities and, if so, at which levels. It is instructive to compare this state of affairs with the phonological awareness literature, which—owing to its status near the top of the Anglophone research agenda—boasts numerous studies of this type, both longitudinal and cross-sectional (e.g., Anthony & Lonigan, 2004; Carroll, Snowling, Hulme, & Stevenson, 2003; Hoien, Lundberg, Stanovich, & Bjallid, 1995; Wagner, Torgesen, & Rashotte, 1994; Yopp, 1988).

Theories regarding the mechanisms underlying fluency are as bountiful as the definitions surveyed above. Wood et al. (2002) speculate on links between their two hypothesized components of fluency and cytoarchitectural (pyramidal/granular, ventral/dorsomedial) and neurochemical (dopaminergic vs. noradrenergic) cortical subsystems, and genetic loci (Chromosome 1 vs. 6). Berninger et al. (2002) propose associations between their tripartite model (speed, automaticity, and executive functioning) and distinct brain loci (cerebellum, striatum, and left frontal region). Breznitz (2006) claims that fluency represents the synchronization of several brain-related processes (principally visual–orthographic and auditory–linguistic), each operating according to different processing and timing constraints.

At the cognitive/behavioral level of explanation, Torgesen (2002) argues that the size of a child’s sight vocabulary is the key determinant of text reading rate and is largely dependent on reading volume. Bear (1991) also sees orthographic knowledge as central to the development of word-level fluency. Berninger et al. (2002) propose that morphological knowledge is critical to the development of word reading fluency, whereas Schreiber (1991) and Dowhower (1991) view syntax as crucial.

The debate over rapid automatized naming (RAN). The controversy over the role of RAN in early reading (see, e.g., Chiappe, Stringer, Siegel, & Stanovich, 2002; Katzir, Kim, Wolf, Morris, & Lovett, in press; Manis & Freedman, 2002; Savage & Frederikson, 2005; Vellutino et al., 2004; Vukovic & Siegel, 2006; Wolf &
Bowers, 1999) also seems to be muddled with accuracy versus speed/fluency issues. Any speeded measure should correlate more strongly with timed or rate-dependent measures of reading than simple untimed accuracy. Because the developmental literature in English is so heavily weighted toward accuracy rather than fluency, with the reverse true of more consistent orthographies, RAN should emerge as a strong predictor of reading (rate/fluency) in regular orthographies. This would partly explain the stronger association between RAN and reading in more regular orthographies—such as German (Wimmer, 1993) and Dutch (de Jong & van der Leij, 1999, 2003)—relative to English, and the weaker association with phonological awareness, which is usually an untimed measure. This “method” account of differential patterns of association can be extended to English. While children are in the (prolonged) learning-the-code phase (up to Grade 3), accuracy is the focus of ability differences; thus, RAN should be a poor predictor of word-recognition performance. After children have acquired the complex spelling–sound code, speed and fluency come to the fore; thus, RAN should be a significant predictor. More generally, if any two variables, such as speed and accuracy, are imperfectly correlated, there will necessarily be “double dissociations” with a “double-deficit” group, a “no-deficit” group, and two smaller “single-deficit” groups. A speeded measure, such as RAN, compared with a nonspeeded measure, is likely to tap speed/accuracy dissociations. It remains to be seen to what extent the classic dual-route distinction between phonological and surface dyslexia, a purely accuracy-based dichotomy, relates to accuracy/speed differences, particularly in the case of more conventional (i.e., consistent) orthographies.

Does Fluency Invariably Flow From Practice?

Not only is there little consensus regarding the nature and definition of fluency but there are several lines of evidence that appear to impeach the belief that fluency and automaticity develop as a natural consequence of practice after word identification accuracy has been attained (Carver, 1990; Chall, 1983; LaBerge & Samuels, 1974; Nathan & Stanovich, 1991; National Reading Panel, 2000). First, “compensated” adult English-speaking dyslexics, many of whom have successful records in advanced academic (college and university) studies, are slow effortful word readers despite high levels of reading accuracy. Clearly, years of constant practice have not proven sufficient for attaining speed and fluency (see, e.g., Bruck, 1990; Lefly & Pennington, 1991; Shaywitz et al., 2003). Childhood dyslexia in consistent orthographies (e.g., German, Italian, Greek, Norwegian, and Hebrew) reveals the same pattern of high accuracy but slow, dysfluent reading (Hagtvet & Lyster, 2003; Porpodas, 2006; Share & Shalev, 2004; Wimmer, 1993). van der Leij and van Daal (1999), for example, found that Dutch childhood dyslexics are slower than normal readers even when accurately reading highly familiar words. Second, in the remediation of reading disability, successful training of phonological awareness and decoding brings accuracy and (word) comprehension gains but not fluency gains (see reviews by Lyon & Moats, 1997; M. S. Meyer & Felton, 1999; Torgesen, 2005). Thaler et al. (2004) have shown that young German dyslexics with high reading accuracy have persistent slowness that is largely resistant to change even after massive practice (see also de Jong & Vrieland, 2004). Finally, Breznitz (2006) has demonstrated that modest individualized acceleration of reading rate causes increases in both reading accuracy and comprehension among dyslexics and normal readers. The fact that Breznitz’s data generalize to normal readers is especially troublesome for the conventional view of speed and fluency as “dependent” variables. Similarly, Lurie and Share (2007) found that, among normal third-grade readers of highly regular pointed Hebrew, the formation of accurate orthographic representations for new words depended on decoding accuracy and decoding fluency (see also Share, 2004). Wood et al. (2002) have also suggested that fluency training need not be delayed until after accuracy is attained.

Viewed within the context of the unfamiliar-to-unfamiliar/ novice-to-expert framework discussed earlier, training phonological skills and decoding accuracy should influence the algorithmic/analytic aspects of initial word learning (deciphering) but may not be sufficient for attaining automatization in fluent skilled reading. Models that regard well-specified representations as the key to fluent orthographically-based word recognition—such as Ehri (1992), Perfetti (1992), and Share (1995)—may be illuminating only one piece of a more complex puzzle in the novice-to-expert transition.

Unresolved Measurement Issues

The problem of oral reading inaccuracy that continues to preoccupy English-language researchers has impeded progress on fluency-related issues in another way. Substantial inaccuracy introduces obstacles in the measurement of speed and latency. When individual item responses are measured, it is conventional to discard reaction times for incorrect responses. Thus, valid reaction times are based on nonidentical sets of stimuli that may differ in uncontrolled ways. One solution would be to examine only those items on which all, or nearly all, subjects perform errorlessly. This, however, might leave a very restricted range of items. Alternatively, the data for incorrect responses can be treated as missing, and these scores can be estimated. Neither solution is without its problems.

As regards list-based assessments of fluency, graded word lists are standard fare for English investigations of reading ability. These lists are typically graded for length and “difficulty,” often beginning with short, high-frequency words that become progressively longer and less common. This is also the format of many popular tests of isolated word recognition (e.g., Wide Ranging Achievement Test, Woodcock–Johnson–III, and Burt Word Reading Test) and text-reading (e.g., Gray Oral Reading Test–4: Woodcock–Johnson–III, and Burt Word Reading Test) and text-reading (e.g., Gray Oral Reading Test–4: Wiederholt & Bryant, 2001; Neale Analysis of Reading Ability—Revised: Neale, 1997). Raw reading rate uncorrected for errors, however, will be confounded with difficulty; easier material is read faster. If fluency is assessed as the number of items correctly read in a given time interval (see Brus & Voeten, 1979), this “words-(correct)-per-minute” measure (assuming a graded list) is likely to differ significantly both within ability levels at a given age and across developmental levels. This dilemma calls for solutions such as maintaining uniformity of item difficulty, repeated “cycles” of difficulty gradation within the list (each cycle consisting of just a few items; see Shatil & Share, 2003), or conversion to a common metric (such as syllables per minute, ems, letters, etc.). Alternatively, reading rates can be examined at predetermined levels of accuracy.
It is also standard practice to measure vocalization \textit{onset} latency in the English-language literature, but, in regular orthographies, a case can be made (see, e.g., Landerl et al., 1997; Naslund & Schneider, 1996; Thaler et al., 2004) for using voice \textit{offset}, especially for inexperienced or disabled readers whose (oral) reading is characterized by laborious sequential decoding (Ziegler & Goswami, 2005). Bowey and Muller (2005) have also argued against voice onset times among unskilled readers of English for similar reasons. In list reading contexts (e.g., Frith et al., 1998), false starts, self-corrections, and hesitations may seriously confound speed differences when accuracy levels are below ceiling. Again, in a regular orthography with high accuracy, measuring list-reading fluency is a manageable problem. In English, however, this problem can be acute.

A child who reads 50 words—all correctly in 1 min—is assigned a fluency score of 50 (words-per-minute or wpm). A second child who reads the same 50 words in 1 min but makes 10 errors is reading at the same (raw) rate but is assigned a score of 40 wpm. The only difference between these two readers, however, is accuracy. Similarly, a child who reads 60 words in 1 min with 10 errors also gains a score of 50 wpm. It is clear that substantial variability in rate and accuracy can yield fluency or efficiency scores that mask important accuracy and rate discrepancies. These considerations, furthermore, apply to any metric that combines accuracy and rate into a single fluency measure.

These complications are not overwhelming in studies involving regular orthographies but may be insurmountable in outlier English. In English, measuring fluency (or "efficiency") by combining accuracy and rate is like conflating height and weight by referring to people as "big" or "little." Like accuracy and speed, weight and height are correlated imperfectly—tall people are usually heavier than short people, but is a tall and thin individual "bigger" than a short and stout individual (who might weigh more)? A growing number of reading disability theorists have called for a separation between rate and accuracy difficulties among disabled readers (Berninger et al., 2002; Lovett, 1987; Wolf & Katzir-Cohen, 2001). Depending on the size of the correlation, two imperfectly correlated measures will, by definition, result in four different profiles: (1) fast and accurate readers (the most proficient), (2) slow and inaccurate readers (the poorest readers)—together making up a majority of the sample if the correlation is greater than zero, (3) fast-inaccurate readers, and (4) slow-accurate readers. The size of the latter two groups (ranging between 0% and 25% of the sample) will be inversely proportional to the magnitude of the speed/accuracy relation.

\textbf{The Role of Reading Accuracy in Definitions of Reading Disability}

Definitions of dyslexia in the English-speaking world have traditionally been operationalized using measures of accuracy, specifically oral reading accuracy. In a special issue of the \textit{Journal of Learning Disabilities} devoted to definitions of learning disability, two of the leading North American dyslexia authorities, Siegel (1999) and Stanovich (1999), proposed that reading disability be assessed on the basis of word and/or pseudoword reading. In both cases, only measures of oral reading \textit{accuracy} were stipulated. As Lyon et al. (2005) have protested, [R]eadling fluency is rarely assessed in current identification procedures . . . Slow-reading students who score within the average range on both the untimed reading measures and the IQ test will typically be denied services because there is no discrepancy—even though they also have a disability that requires specialized services and/or accommodations. (p. 267)

The British Psychological Society (1999) was the first English-speaking organization to include fluency in the definition of dyslexia (see also Lyon, Shaywitz, & Shaywitz, 2003). Interestingly, the British borrowed their definition from the Dutch (Gersons-Wolfenberger & Ruijsenaars, 1997), which—being a more regular orthography—emphasizes fluency and automaticity rather than accuracy.

Timed tests of word and pseudoword reading speed and accuracy are standard practice in non-English dyslexia research (e.g., in German: Wimmer, Mayringer, & Raberger, 1999); in French: Bosse, Tainturier, & Valdois, 2007; in Norwegian: Fagerheim et al., 1999; in Italian: Facetti et al., 2006; in Hebrew: Breznitz, 1997). The preference for accuracy-based definitions in English and fluency-based definitions among non-English researchers creates additional barriers for cross-linguistic comparisons in this field.

\textbf{English Spelling–Sound Ambiguity and the \textit{Wait-to-Fail} Model of Reading Disability}

English spelling–sound ambiguity also plays a role in the fiercely debated IQ-discrepancy definition of dyslexia. This definition is founded on the gap between attainment in reading and aptitude as measured by IQ tests (Lyon et al., 2005; Share, McGee, McKenzie, Williams, & Silva, 1987; Siegel, 1989; Stanovich, 1991; Stuebing et al., 2002). As already noted, the complexity of English spelling–sound relations leads to a considerably longer period of code learning than occurs in consistent orthographies, extending over about 3 years of schooling (Chall, 1983/1996; Harlaar, Dale, & Plomin, 2007; Seymour et al., 2003; Singer, 1978). It is probably not until the fourth grade that most children acquire the requisite word recognition skills to enable them to successfully read any material they can comprehend through listening. In Jeanne Chall’s (1983/1996) terms, it is at this point that most students progress from "learning to read" to "reading to learn" (Shankweiler & Fowler, 2004, p. 498). Many educational practices have been founded on the assumption of a 3-year learning-to-read phase. The establishment of remedial reading centers in Britain to which children are not referred until 9 years of age was predicated on the belief that it was premature to diagnose and remediate reading failure before reading (primarily phonics) instruction had run its course (Collins, 1961). Similarly, the IQ-discrepancy model, still enshrined in U.S. Federal law, requires a severe discrepancy that is not normally evident until around 3 years after children enter school (Fletcher et al., 2002; Lyon et al., 2005). This leads to what Lyon et al. (2005) term a \textit{wait-to-fail} model: “Measurement practices . . . are biased against the identification of children before age nine” (p. 263). “Because achievement failure sufficient to produce a discrepancy from IQ cannot be reliably measured until a child reaches approximately 9 years of age, the use of IQ-discrepancy constitutes a ‘wait-to-fail’ model”
(Lyon et al., 2005, p. 266) militating against early identification and intervention. The alignment of the 3-year delay with the Anglophone learning-to-read triennium, a product largely of English spelling–sound complexity, may not be coincidental. Sadly, the efficacy of remediation is severely curtailed after nine years of age (Lyon et al., 2005).

**Oral Decoding**

The goal of word reading is access to meaning (Rayner et al., 2001). Silent understanding rather than oral reading is the literacy benchmark in knowledge-based societies (Lundberg, 1994; Torgesen, 2005). Of course, oral recitation has an essential role in both secular and religious ceremony and in reading aloud to preliterate children. Nonetheless, both ontogenetically (Edfeldt, 1960; Kragler, 1995; Levin, 1979; Wright, Sherman, & Jones, 2004) and historically (Manguel, 1996; McLuhan, 1962; Saenger, 1982; Venezky, 1984), reading aloud (typically in dyadic or communal settings) is superseded by the more efficient mode of silent and private reading.

English-language reading research, however, is dominated by measures of oral rather than silent reading. Computational modeling of skilled word reading focuses almost exclusively on reading aloud (M. Coltheart et al., 1993, 2001; Perry et al., 2007; Plaut et al., 1996; Seidenberg & McClelland, 1989; Zorzi et al., 1998). Tests of the ability to read aloud a graded list of words (or pseudowords) are benchmark measures of English word identification. Oral word-reading instruments are the primary criteria recommended by leading reading disability theorists for the definition and diagnosis of reading disability (British Psychological Society, 1999; Siegel, 1999; Stanovich, 1999). Oral miscue analysis (Goodman, 1973) and New Zealand’s running records (Clay, 1979), all oral measures, are standard assessment tools in their respective traditions. Ironically, even the contemporary English-language literature on reading disability subtypes is based almost exclusively on the naming of various types of items. Unsurprisingly, these are pseudowords and exception words (see, e.g., Castles & Coltheart, 1993; Harm & Seidenberg, 1999; Manis et al., 1996; Stanovich et al., 1997). Remarkably, some of the leading theoretical work on the development of reading comprehension skill has been based on an oral reading measure: the Neale Analysis of Reading Ability (Nation, 2005; Oakhill, 1993).

This partiality toward oral reading in English-language reading research is not surprising given the magnitude of the word identification problems posed by English spelling–sound ambiguity: The simplest way of verifying the accuracy of word identification is, of course, reading aloud. I am not suggesting that the bias toward oral reading is solely the product of English spelling–sound ambiguity; oral measures have the virtue of simplicity, and precision in oral recitation has, historically, been accorded greater importance than understanding. Nevertheless, it seems reasonable to assume that the unique word identification problems in English have predisposed English-language reading science toward pronunciation accuracy as the yardstick of word identification.

Reading aloud, however, does not necessarily involve access to meaning—the goal of word reading (see, e.g., Baron, 1977; M. Coltheart, 1978). Ironically, the two cofounders of the regularity-oriented dual-route approach (Coltheart and Baron) initially dis-qualified the very task that has served as the mainstay in developing and testing the dual-route model and its subsequent competitors. Writing in 1978, M. Coltheart observed that

A wide variety of information-processing tasks has been used over the past decade for investigating this [lexical access] and other aspects of reading. The suitability of many of these tasks for investigating lexical access can, however, be questioned. Some tasks can be criticized because they do not necessarily involve lexical access . . . Examples of the first kind of task are . . . reading aloud. (p. 170)

Baron (1977) also pointed out that

There are certain tasks that clearly do not qualify as measures of extracting meaning. Among these tasks are the pronunciation tasks I have been discussing up to now . . . and lexical decision tasks . . . While these tasks might involve the use of meaning, they do not require it. (p. 180)

Rayner and Pollatsek (1989) expressed similar doubts about the validity of oral text-reading measures. They drew attention to the fact that the eye is typically several words ahead of the voice:

It is entirely likely that the words produced by the speech mechanism in oral reading are strongly influenced by processes occurring after the lexicon has been accessed, and thus the resulting data may say little about how the lexicon is initially contacted to arrive at the meaning of a given word. (p. 181)

Another concern is that the pronunciation of the written word appears to depend on exhaustive and well-specified phonological representations of the type that may not fully correspond to the phonological representations required for nonoral word recognition and meaning access (see Berent & Perfetti, 1995; Frost, 1998). Oral reading rates are typically slower than silent reading rates (Barker, Torgesen, & Wagner, 1992; Carver, 1990, chapter 14; de Jong & Share, 2007; Levin, 1979), consistent with the idea that silent reading involves less exhaustive phonological processing than oral reading.

The possibility that oral and nonoral reading rely on representations that differ in completeness is bolstered by the fact that representations in written language are usually the maximally distinct, “careful” or “formal” forms, as opposed to the more abbreviated “informal” or “vernacular” forms that typify running speech (Elbro, 1996; Fromkin & Rodman, 1974; Lundberg, 1994). A correct “careful” pronunciation of the printed word and has three phonemes, but in day-to-day speech, the final /ə/ is often elided (as are most word-final stops), and the initial vowel is reduced to a schwa (some authors represent this colloquial form with the spelling ’n). This phonological “reduction” applies to most words (although the term “augmentation” might be more apt if we consider that, developmentally, speech forms precede written forms). Thus, the “correct” pronunciation of the printed word, especially when presented in isolation, frequently diverges from the reduced (or “unaugmented”), but more familiar, vernacular form. It is, therefore, a misnomer to speak of the pronunciation of a word; word pronunciation is a family of spoken forms that varies according to pragmatics, speech rate, and dialect. In this sense,
learning to read is also learning new speech variants. In view of the phonological distance between spoken and written forms, oral reading, especially of isolated words, may depend on knowledge of the formal (literate) augmented form, whereas the everyday unaugmented form may be sufficient for meaning access during silent reading.

The foregoing considerations suggest that oral and nonoral reading modes have much in common but are unlikely to invoke identical cognitive processes. If so, then the dependence on oral reading in current reading research may yield an incomplete picture of the nature of word reading. Unfortunately, little empirical data bear on this issue. Numerous studies in the earlier part of the 20th century (for reviews, see Carver, 1990; chapter 4) directly compared oral and silent (text) reading, but postwar cognitive psychology (with rare exceptions, such as Barker et al., 1992; Juel & Holmes, 1981) lost interest in this issue.

Although very few studies have directly compared oral and silent reading, indirect evidence converges on the conclusion that oral and nonoral modes have both common and unique features. To the extent that naming and lexical decision data tap the oral/nonoral issue, studies have repeatedly demonstrated consistent differences between naming and lexical decision as regards the influence of phonology. Regularity effects, at least for low-frequency words, are ubiquitous in naming (e.g., Andrews, 1982; Baron & Strawson, 1976; Weekes, 1997) but not in lexical decision (e.g., Seidenberg & McClelland, 1989; Seidenberg, Waters, Barnes, & Tannenhaus, 1984). These effects are not idiosyncratic to English but apply to regular orthographies as well. For example, in Hebrew, word pointing (the addition of vowel diacritics) and phonological ambiguity affect latencies in naming but not in lexical decision (Bentin, Bargai, & Katz, 1984; Frost, 1995; Koriat, 1984). Length effects, especially for pseudowords, are found in naming but not in lexical decision tasks (Ferrand & New, 2003; Frederiksen & Kroll, 1976; Valdois et al., 2006; Weekes, 1997). Frequency effects, in contrast, are stronger in lexical decision than in naming (e.g., Balota & Chumbley, 1984; Forster & Chambers, 1973; Frederiksen & Kroll, 1976; McCann & Besner, 1987; Scarborough, Cortese, & Scarborough, 1977).

Many studies have demonstrated pervasive effects of phonology, even in silent reading tasks and among skilled adult readers (see, e.g., V. Coltheart, Patterson, & Leahy, 1994; Folk, 1999; Frost, 1998; Miellet & Sparrow, 2004; Perfetti, Bell, & Delaney, 1988; Perfetti & McCutchen, 1982; Pollatsek, Rayner, & Lee, 2000; Tseng, Hung, & Wang, 1977; Van Orden, 1987; but see Van Orden & Kloos, 2005); however, phonology has a reduced role in silent reading (or at least in lexical decision) relative to oral reading. Given the reservations expressed by many theorists as to the ecological validity of both naming and lexical decision tasks (e.g., Baron, 1977; M. Coltheart, 1978; Rayner et al., 2001; Rayner & Pollatsek, 1989; Seidenberg, 1985), it is reassuring that the pattern of phonological attenuation in silent lexical decision concurs with findings from studies examining the reading of connected text. For example, Juel and Holmes (1981) found stronger effects of regularity and syllable length in oral sentence reading compared with silent reading among children, particularly among poor readers. Similarly, Barker et al. (1992) found that both oral pseudoword repetition and phoneme deletion correlated more strongly with oral text reading rate than with silent text reading rate among third graders.

Consistent with the natural variability of phonological forms in general, and the phonological distance between speech and writing in particular, the experimental data suggest that overt naming involves more thorough phonological analysis and/or more completely specified representations than nonoral word identification. Moreover, overt naming may involve less attention to orthographic structure or to meaning than silent reading (Corcos & Willows, 1993).

Another line of evidence indicating that overt pronunciation may provide only a monochrome picture of the word reading process comes from case studies of skilled readers with relatively poor word and pseudoword naming (e.g., Campbell & Butterworth, 1985; Holmes, 1996; Stothard, Snowling, & Hulme, 1996). Holmes (1996), for example, reported the case of KQ, a highly literate university student with very poor oral reading of words and pseudowords but whose performance in lexical decision and spelling choice tasks was virtually perfect. Lexical decision times for the very same pseudowords that KQ was unable to read aloud were faster than any control subject, and silent reading comprehension was excellent. It is worth remarking that such “discrepant” readers are not particularly rare (see Jackson & Doellinger, 2002; Shankweiler et al., 1999). Furthermore, in many of these cases, proficient (silent) reading comprehension cannot be attributed to superior verbal ability or verbal memory (Jackson & Doellinger, 2002).

Studies of individuals with severe speech and/or hearing impairments also affirm the separability of oral and nonoral modes of reading. The congenitally deaf population as a whole has generally low levels of reading comprehension skill (Conrad, 1979; Wauters, van Bon, & Tellings, 2006); however, a substantial minority of profoundly deaf readers possesses excellent silent word recognition skills despite nonexistent oral decoding (e.g., Burden & Campbell, 1994; Hanson, 1986; Miller, 2004; Wauters et al., 2006). In the case of selective speech impairment, Foley and Pollatsek (1999) found that substantial numbers of dysarthric and anarthric individuals scored in the normal range on silent reading comprehension (see also Cossu, 1999). Patterson, Graham, and Hodges (1994) have also reported cases of dementia patients with poor performance on tasks such as picture naming and category fluency, but who were only slightly impaired on tasks requiring no speech output, such as word-picture matching.

Yet another “special” group, dyslexics are well known for their relative disassociation between (silent) orthographic strengths and (oral) phonological weaknesses (e.g., Rack, Snowling, & Olson, 1992; Share, 1995; Siegel, Share, & Geva, 1995; Stanovich & Siegel, 1994). Dyslexics typically demonstrate deficits relative to reading-age controls in pseudoword reading, but they perform comparably or better than controls on orthographic processing.

8 Extreme cases of divergence between spoken and written phonological forms (i.e., “digglossia”)—such as Arabic (Ferguson, 1959) or Kannada (Karathan, 2006)—merely represent the extreme point on a spoken/written continuum. The fact that oral reading accuracy is uncorrelated with reading comprehension in Arabic clearly underlines this divergence (Abu-Rabia, 2001; Saigh-Haddad, 2003; see also R. C. Anderson et al., 2003).

9 The “phonological processes” used by normal infants, as yet unable to produce adult speech, to simplify adult speech (Ingram, 1976; Khan & Lewis, 1986) and also by adults in informal conversational speech may be a fruitful avenue for the investigation of oral/silent differences in phonology.
Some dyslexics are even able to attain orthographic skills comparable with chronological age-matched normal readers in a second language (Bekebrede, van der Leij, & Share, in press; van der Leij & Morfidi, 2006). In addition, Stanovich and Siegel (1994) found that dyslexics have more severe difficulties with pseudoword tasks requiring overt pronunciation than with pseudoword tasks requiring only implicit decoding, such as R. K. Olson et al.’s (1985) phonological judgment task (see also Leinonen et al., 2001). Similarly, Share, Jorm, Matthews, and Maclean (1988) found that naming was relatively easier than lexical decision for normal second-grade readers but not for disabled readers.

The oral/nonoral dissociations discussed above have all involved cases in which oral reading is more severely impaired than nonoral reading. However, there is also evidence pointing to true double dissociation. In a study of normally developing third graders learning Hebrew, Canaan and Share (2007) used a semantic categorization task to assess silent word reading (children were required to circle all pseudohomophones that named foods in a list of 100 items) and a naming task to assess oral reading (reading aloud these same pseudohomophones). The correlation between naming times (accuracy was at ceiling levels) and categorization was .61, indicating that less than half of the variance was shared. A similar correlation coefficient was obtained when the same list of words used in the semantic judgment task, but conventionally spelled, was read aloud (Share, unpublished data, 2006). The oral task, furthermore, was more strongly correlated with phonological abilities than the silent task, whereas the silent task was more closely tied to orthographic and morphological knowledge. A subgroup of poor (oral) decoders who were good (silent) word readers were characterized by good semantic, orthographic, and morphological knowledge but poor phonological skills. A second group with normal pseudoword decoding, but poor silent word reading, revealed the inverse profile: good phonological skills, but poor semantic, orthographic, and morphological knowledge. Similarly, Bekebrede et al. (in press) partitioned a group of Dutch adolescent dyslexics into two subgroups on the basis of (L2) English orthographic knowledge and found that the subgroup with superior orthographic skills (both groups displayed poor phonological skills) was significantly better at silent (Dutch) word identification than the other subgroup, but the subgroups did not differ on oral word reading fluency.

Yet another line of evidence highlighting oral/silent disparities derives from neuropsychological and neuroimaging evidence suggesting that the primary substrate of sound-based representations of speech (the superior temporal cortex) interfaces in task-dependent ways with other systems (Poeppel, 2001). Poeppel (2001) summarized several articles indicating that the auditory–motor/articulatory interface and the lexical–semantic interface occupy distinct brain pathways.

Levin’s (1979) review sums up the oral/silent issue succinctly: “There are enough similarities [between oral and silent reading] to warrant attention to oral reading, although the two modes are different enough to suggest caution in extrapolating from one to the other” (pp. 37–38). It follows that a research literature skewed toward oral reading may provide an incomplete picture of the reading process. The English-language focus on oral reading, like accuracy, may overestimate the role of phonological variables in word reading, underestimate the role of other nonphonological factors, and deflect research attention from less tractable, yet more crucial, issues (such as silent reading and meaning access).

Phonological Awareness

Awareness of the phonological units represented by a given writing system is positively correlated with reading ability in every language studied to date (National Reading Panel, 2000). This affirms a universal phonological axiom regarding writing; orthographies must represent units of spoken language. Whether these units are syllables, morae, subsyllabic units, or phonemes will depend on the particular language (Goswami, 1999; Mattingly, 1985).

The awareness of phonemes and the acquisition of alphabetic literacy does not appear to be an integral part of humans’ biological preparedness for rapid early spoken language acquisition, but, in the words of Bowey (2005), is “inextricably linked” to learning to read an alphabetic orthography (p. 168). Moreover, as an integral component of alphabetic literacy learning, phonemic awareness is closely tied to the particular orthography concerned, the nature of the units represented, and the fidelity of the spelling–sound mapping. Thus, the lack of consistency in English should exacerbate difficulties in acquiring phonemic awareness. It is now clear that the scale of this problem in English does not generalize to more regular orthographies and that the unique role of phonemic awareness in English has hindered consideration of more universal aspects of phonology that apply across orthographies. Phonic awareness is a core component of learning to read in every possible alphabetic orthography; however, the extreme degree of nontransparency in English has exaggerated the role that phonemic awareness plays in more conventional alphabets and has overshadowed issues that have critical importance across orthographies.

Phonemic Awareness and Orthographic Transparency

Several lines of evidence converge to indicate that phonemic awareness constitutes an integral part of alphabetic reading acquisition dependent on the experience of learning to read an alphabetic script (see, e.g., Bertelson & de Gelder, 1989; Burgess & Lonigan, 1998; Castles & Coltheart, 2004; de Santos Loureiro et al., 2004; Ehri, 1979; Holopainen, Ahonen, & Lyytinen, 2001; I. Y. Liberman, Shankweiler, & Liberman, 1989; K. Lukatela et al., 1995; Morais, Alegria, & Content, 1987; Perfetti, Beck, Bell, & Hughes, 1987; Shankweiler & Fowler, 2004) and on the architecture and transparency of that specific script (e.g., Burgess & Lonigan, 1998; K. Lukatela et al., 1995; Mann & Wimmer, 2002; Share & Blum, 2005; Tolchinsky & Tebiski, 1998). A more detailed exposition of the script-dependent claim can be found in Share (1995), but two points are pertinent here. First, several studies have shown that both adult illiterates who are intellectually and linguistically competent (Bertelson, de Gelder, Tfouni, & Morais, 1989; de Santos Loureiro et al., 2004; K. Lukatela et al., 1995; Morais, Cary, Alegria, & Bertelson, 1979; Scholes & Willis, 1987) and individuals literate only in nonalphabetic scripts (Mann, 1986; Prakash, 2003; Read, Zhang, Nie, & Ding, 1986) have poor
awareness of phonemes. This demonstrates that sensitivity to phonemic segments does not arise spontaneously outside the context of learning to read an alphabetic script (Bertelson et al., 1989; Morais & Kolinsky, 2005). Second, the largest performance gains on tests of phonemic awareness occur during the 1st year of alphabetic reading instruction, largely irrespective of age (e.g., Barron, 1991; Bentin, Hammer, & Cahan, 1991; Calfee, Lindamood, & Lindamood, 1973; Cardoso-Martins, 1991; Carrillo, 1994; de Jong & van der Leij, 1999; Duncan, Cole, Seymour, & Magnan, 2006; Hanley, Tzeng, & Huang, 1999; I. Y. Liberman, Shankweiler, Fischer, & Carter, 1974; Mann & Wimmer, 2002; Morais, Content, Bertelson, Cary, & Kolinsky, 1988; Torgesen et al., 1989). Most preliterate preschoolers are unable to access phonemes (Bowey, 2000, 2005; Carrillo, 1994; Cossu, Shankweiler, Liberman, Katz, & Tola, 1988; I. Y. Liberman et al., 1974; Oney & Durgunoglu, 1997; Share & Blum, 2005; Share & Gur, 1999). Method of instruction is also influential. Code-emphasis instruction accelerates the development of phonemic awareness to a greater extent than whole-word instruction (e.g., Adams, 1990; Alegría, Pignot, & Morais, 1982; Caravolas & Bruck, 1993; Carrillo, 1994; McBride-Chang, Bialystok, Chong, & Li, 2004; Perfetti et al., 1987).

Collectively, the data from both illiterate and preliterate samples indicate that most children develop an awareness of phonemes as they learn about letters and the alphabetic mapping principle in the course of formal reading instruction. The influence of the particular alphabetic code is also apparent in that English speakers will often insist that there are more sounds in the word *pitch* than the word *rich* (Ehri & Wilce, 1983; Freiman & Kessler, 2003; Tunnner & Nesda, 1985). Conversely, Scholes (1995) reported that university undergraduates are more successful at phoneme deletion when sounds are represented unambiguously in spelling (e.g., /raft/ without /f/ is /raft/) than when sounds are not uniquely represented by a single letter (e.g., *mixed* without /k/ = *mist*). “[L]earning to read is learning to hear speech in a new way” (D. R. Olson, 1994, p. 85).

In summary, the available evidence indicates that phonemic awareness is best categorized as a reading subskill, not as a universal and emergent linguistic capability (Share, 1995). An alphabetic script is, first and foremost (although not solely), a phonemic code or “blueprint”—a graphic dissection of the segmental structure of spoken words. Phonemic awareness and letter knowledge have been labeled corequisites to alphabetic literacy (Share, 1995) or codeterminants (Bowey, 2005). Ehri and Soffer’s (1999) term *graphophonemic awareness* aptly captures the inseparability of the two. Attempts to disentangle phoneme awareness and orthographic knowledge in the quest for ever “cleaner” studies of phonemic awareness unconfounded by letter knowledge are consequently misconceived.

If we accept the view that phonemic awareness is script dependent, then it follows that the clarity of this phonemic blueprint should affect the development of phonemic awareness. A transparent code with consistent one-to-one letter-sound relations should promote access to phonemes, whereas opacity should impede it. The data generally support this prediction.

In phonemically transparent orthographies, phonemic awareness is acquired more rapidly and easily than in less transparent orthographies (e.g., in German: Mann & Wimmer, 2002; in Dutch: de Jong & van der Leij, 1999; in Hungarian: Csepe, 2006; in Czech: Caravolas & Bruck, 1993; in Italian: Cossu et al., 1988; in Welsh: Spencer & Hanley, 2003; in Turkish: Oney & Durgunoglu, 1997; in Finnish: Holopainen et al., 2001; in Greek: Nikolopoulos et al., 2006; and in Hebrew: Geva et al., 1993). For example, in their study of phonemic awareness among first-grade Canadian English–Hebrew bilinguals, Geva et al. (1993) obtained a correlation of .62 for English but only .32 for Hebrew (correlations for pseudoword reading and phonemic awareness reproduced this disparity: .68 vs. .39, respectively). Almost identical figures were found in two longitudinal investigations with large monolingual samples of English-speaking (Australian) and Hebrew-speaking (Israeli) children (Share et al., 1984; Shatil & Share, 2003). To the extent that pseudoword spelling reflects phonemic analysis, several cross-linguistic studies of spelling acquisition (see Caravolas, 2006, for a review) provide additional support for an attenuated reading-phonemic awareness relation in more transparent orthographies. Even dyslexics are relatively successful at phoneme manipulation in regular orthographies, at least according to accuracy criteria (e.g., de Jong & van der Leij, 1999; Wesseling & Reitsma, 2000; Wimmer, 1993; Wimmer & Goswami, 1994).

Concerns regarding the perils of cross-language comparisons, such as item and subject incommensurability, cannot be dismissed entirely, as discussed earlier in the article. Again, however, fewer studies that have compared well-matched subjects from the same culture (e.g., Spencer & Hanley, 2003) or different languages in the same sample of bilingual children (Geva et al., 1993) have confirmed a weaker reading–phonemic awareness relation in orthographies that are more transparent than English.

**Functional Opacity and the Timing Hypothesis**

The strength of the phonemic awareness–reading association appears to depend on when (and how) these abilities are assessed (see de Jong & van der Leij, 1999, 2002; Share & Levin, 1999; Wimmer, 1993) and on the functional opacity of the spelling–sound code for readers of a given developmental level. In a study of Hebrew phonemic awareness, Bentin and Leshem (1993) found a correlation of .33 between phonemic awareness and Hebrew decoding at the *end* of Grade 1. This figure converges with the coefficients cited above for voweled Hebrew at the same developmental point by Geva et al. (1993) and Shatil and Share (2003). Significantly, the correlation was .55 *midway* through Grade 1, when most children have yet to be introduced to all the Hebrew letter–sound correspondences. This is very similar to the figure typically reported at the end of Grade 1 for English. Bentin and Leshem’s data suggest that phonemic awareness is maximally influential when code learning is incomplete; that is, even a transparent script is functionally opaque when a child has learned some, but not all, of the spelling–sound mappings.

In a longitudinal study of the relation between phonemic awareness and word reading accuracy in Latvian—another highly regular orthography—Sprugevica, Paunina, and Hoiien (2006) found

---

11 Orthographic transparency is, of course, not the only factor accounting for cross-linguistic differences. Other factors include phonological structure, timing (stress-timed vs. syllable-timed), and number of syllables (see, e.g., Caravolas & Bruck, 1993; Cheung, Chen, Lai, Wong, & Hills, 2001; Cossu et al., 1988; Duncan et al., 2006; Huang & Hanley, 1995; Saiegh-Haddad, 2003; Share & Blum, 2005).
that a composite measure of phoneme segmentation and phoneme deletion accounted for 27% of the variance in reading accuracy in December of first grade, 9% at the end of Grade 1, but nothing by mid-Grade 2. In a follow-up regression analysis controlling for kindergarten word reading (the autoregressive effect), kindergarten phoneme awareness accounted for significant unique variance in reading in the middle, but not at the end, of Grade 1. The same developmental pattern seen in the Hebrew and Latvian data has also emerged in another two highly regular orthographies: Turkish (Oney & Durgunoglu, 1997) and Finnish (Leppanen et al., 2006).

In Dutch, an orthography characterized by intermediate transparency (Borgwaldt et al., 2005; Bosman, de Graaff, & Gijzel, 2006), de Jong and van der Leij (1999, 2002) found that phonemic awareness—reading correlations declined after the end of Grade 1. They concluded that, unlike English, “the importance of phonological abilities for learning to read appeared to be limited to the 1st year of reading instruction” (p. 467). Similarly, Wesseling and Retisma (2000) found that Dutch phoneme awareness (assessed using a blending task) was a stronger predictor early in Grade 1, when children had been taught around three quarters of the simple grapheme–phoneme correspondences, than at the end of Grade 1, when letter–sound knowledge approached ceiling.

The developmental decline in correlations between phonemic awareness and reading is a consistent finding in transparent orthographies and contrasts sharply with the correlations in English, which remain robust across age (Calfee et al., 1973; Deacon & Kirby, 2004; Rosner & Simon, 1971; Swanson, Trainin, Necochea, & Hammill, 2003; Wagner et al., 1997; but see Kirby, Parrila, & Pfeiffer, 2003). Even English-speaking adults with reading disability have poor awareness of phonemes (e.g., Bruck, 1992; Chiappe et al., 2002; Felton, Naylor, & Wood, 1990). This developmental invariance in English suggests that English remains functionally opaque at all levels of reading ability.

Although some studies (Nikolopoulos et al., 2006; Spencer & Hanley, 2003) have reported strong phonemic awareness—reading correlations beyond Grade 1 in highly regular orthographies, the preponderance of evidence favors the functional opacity hypothesis. If correct, this account of the phonemic awareness—reading relation leads to a novel hypothesis: The primary function of phonemic awareness may be to clean up “noisy” decoding approximations by providing well-specified phonemic templates that make it possible to bridge the gap between imperfect decoding attempts and target pronunciations. This “clean-up” or “filter” hypothesis emphasizes a probabilistic, as opposed to deterministic, view of spelling-to-sound conversion consistent with the earlier discussion of phonological form as a family of speech variants.12

In summary, phonemic awareness is likely to be equally important in consistent and inconsistent orthographies but at different phases in development. The normally brief developmental timescale in transparent orthographies is unusually protracted in English, perhaps indefinitely. Thus, the importance of phonemic awareness is inflated relative to its more time-limited role in conventional orthographies.

Minimalist Versus Maximalist Views of Phonology

Contemporary discussion of the role of phonology in reading acquisition is dominated by what might be termed a narrow or “minimalist” view of phonology in which “phonology” is interpreted as synonymous with phonological awareness (see, e.g., Wolf & Bowers, 1999). Substantial variance in reading ability, however, is attributable to a range of difficulties in processing speech-based information. These difficulties pervade perception, learning, and memory and are manifest in a wide variety of tasks (for reviews, see Beaton, 2004; Brady, 1997; Elbro, 1996; Frith, 1997; Jorm & Share, 1983; Kamhi & Catts, 1989; Shankweiler & Liberman, 1989; Share, 1995; Shaywitz, 1996; Snowling, 2000; Stanovich, 1988; Vellutino et al., 2004; Wagner & Torgesen, 1987). “Although problems with phonemic awareness are a hallmark of reading disability, they are only one manifestation of broader deficits in processing phonological information that are typically seen in poor readers” (Shankweiler & Fowler, 2004, p. 503).

Basic phonological processing abilities, unlike phonemic awareness, are inherent in humans’ prewired language learning capacities. Although these abilities are by no means immutable, they all develop prior to the onset of literacy, even from earliest infancy (see, e.g., Boets, Woiters, van Wieringer, & Ghesquiere, 2007; Guttorp, Lepannen, Richardson, & Lyytinen, 2001; Molfeze, 2000; van Leeuwen et al., 2006). These longitudinal/predictive findings furnish crucial evidence for causality (although, of course, less decisive than experimental training data) and critical information for educators seeking early (intrinsic) warning signs that do not primarily originate in environmental deprivation.

The ongoing controversy over the causal status of phonemic awareness (Bertelson, 1993; Castles & Coltheart, 2004; Coles, 2000; Cossu, 1999; Hulme, Snowling, Caravolas, & Carroll, 2005; Scholes, 1998; Stuart, 1998; Ziegler & Goswami, 2005), which has monopolized so much of the phonology debate, has weakened and divided the field. If phonology is equated with phonemic awareness alone, and the latter is coeval with alphabetic literacy learning, then the well-known phonological deficit explanation of reading difficulties becomes vulnerable to charges of circularity (see, e.g., Nicolson, 1996; Tonnessen, 1997). Furthermore, the script-dependent and teaching-dependent nature of phonemic awareness has helped tie this literature to the debate over phonics instruction and thus embroiled the field in the vitriolic reading wars (Coles, 2000; R. Meyer, 2002; Scholes, 1998; Stuart, 1998). Basic phonological processes, in contrast to phonemic awareness, are far less vulnerable to these charges. Moreover, the case for phonology on both sides of the literacy divide (i.e., before and after the onset of schooling) and the significance of the crucial experimental training studies in phonemic awareness are considerably bolstered by acknowledging the connection between basic (preliterate) phonological activities and (peri-literate) phonemic awareness (Baddeley, 1987).

---

12 This applies even to regular words as pointed out several decades ago by the Haskins group (I. Y. Liberman, 1973) in their discussion of the difficulties decoding and blending even simple CVC letter strings, such as bag (see also M. Coltheart, 1978). To arrive at the monosyllable /bæg/, after sounding out three elements, /b/ə, /æ/, and /g/ə, the reader must first perform phoneme elision—the irrelevant but unavoidable schwa sound that accompanies the initial and final stops /b/ and /g/ prior to blending. For an illustration of this problem, see K. Lukatela et al.’s (1995) report of semiliterate Serbo-Croatian women who could sound out words of up to three letters but who were unable to blend the elements into an integrated pronunciation.
A second casualty of an overly Anglocentric perspective on phonemic awareness is the almost exclusive focus on accuracy, yet another factor serving to deflect attention from the issue of speed and fluency in reading development. With rare exceptions (e.g., Good, Simmons, & Kame’enui, 2001; Patel et al., 2004), assessment of phonemic awareness, much like the Anglophone assessment of decoding and reading ability, has been almost exclusively accuracy oriented—the debate preoccupied with issues of task validity, unit size, and timing. Thus, the English phonemic awareness literature, too, has had a share in promoting an accuracy-fixed literature.

The emphasis on phonemic awareness can be regarded as another manifestation of the extreme decipherability problem of English orthography. An orthography that poses extraordinary challenges for the novice will lead to a reading research agenda focused on acquisition-related issues, as discussed above. Many issues relating to the transition to expert performance will be neglected.

Additional Anglocentricities in Reading Science

In the remaining section, I briefly survey several issues that have been focal in English-language reading research but have questionable relevance for more regular orthographies or have been overlooked entirely.

Stage Models of Reading Development

Models of word reading acquisition developed by English-language researchers almost invariably include one or more phases, or stages, in which the novice reader is unable to exploit all the grapheme–phoneme (or higher order orthography–phonology) information available in a printed word, relying instead either on partial letter–sound cues (often in conjunction with contextual cues) or on purely global visual information, such as word length and envelope, or salient visual (nonphonological) features of selected letters. The best-known terms for these stages or strategies are partial alphabetic (Ehri, 1995) and logographic (Frith, 1985), respectively. One or both appear in almost every English-language model of early reading development (e.g., Ehri, 1995; Frith, 1985; Gough & Hillinger, 1980; Harris & Coltheart, 1986; Marsh, Friedman, Welch, & Desberg, 1981; Morton, 1989; Rack, Hulme, Snowling, & Wightman, 1994; Stuart & Coltheart, 1988). Doubts, however, have been raised about the applicability of these stages to other languages and orthographies.

Landerl (2000) observed evidence for partial alphabetic strategies (twenty for twelve) among English first graders but not among matched German readers. In a French longitudinal study from early kindergarten to the end of Grade 1, Spренger-Charolles and Bonnet (1996) found no trace, at any of the four time points examined, of logographic reading, operationalized as reliance on global word form (i.e., length), nonsequentiality of processing, and the use of salient visual cues. Wimmer and Hummer (1990) operationalized logographic reading in their study of German-speaking children in Austria as (a) a failure to read pseudowords despite comparatively successful reading of familiar words and (b) nonattempts when reading unfamiliar words or production of real words visually similar to the target string (lexicalizations). In their comparison of specifically reading-delayed and normally developing first-grade readers who had received 8 months of reading instruction, Wimmer and Hummer found little evidence for logographic strategies. Both groups relied heavily on alphabetic reading strategies. Wimmer and Hummer were, however, able to induce partial alphabetic reading by presenting words briefly for 1 s. Under these conditions, responses revealed partial decoding of initial letters followed by guessing after the word was removed from view. The authors did not discount the possibility that some German-speaking children may identify certain words, chiefly nonalphabetic logos, logographically prior to school entry. Wimmer and Hummer also noted that the Austrian preschool curriculum explicitly discourages exposure to printed words and letter learning, whereas children are immersed in a highly transparent orthography taught via phonics methods at school. Austrian children appear to progress rapidly from nonreading to fully fledged alphabetic reading with little need to develop strategies of the logographic/partial-alphabetic variety (see also Mann & Wimmer, 2002). Similar conclusions have been drawn from studies conducted in German (Mannhaupt, Jansen, & Marx, 1997), in Italian (Job & Reda, 1996, as cited in Job, Peressotti, & Mulatti, 2006), in Greek (Porphodas, 2006), and in Kannada (Karanth, 2006).

Logographic and partial alphabetic stages appear to be largely an English-language peculiarity, a product of an unusually protracted period of early reading development jointly attributable to encouragement of early literacy during the preschool years followed by a prolonged period of code learning. However, both logographic and partial-alphabetic phenomena have also been reported in regular orthographies (Cardoso-Martins, 2001; Share & Gur, 1999). Cardoso-Martins (2001) found evidence of partial alphabetic reading among beginning readers of Portuguese taught via a whole-word method but not among a comparable group receiving phonics instruction. Share and Gur (1999) found evidence of logographic and partial-alphabetic (“phonetic-cue”) reading among Israeli kindergarten children who had not yet been exposed to formal reading instruction, but no evidence of either partial alphabetic or logographic reading has been reported among first graders in Israel.13

The functional opacity argument developed earlier in regard to the phonemic awareness–reading relation can be applied to these findings as well. When children begin reading in Grade 1, or in a preformal context, they have an incomplete mastery of the spelling–sound system, owing either to an opaque orthography or to teaching methods that make the orthography functionally opaque. In these situations, transitional phenomena, such as logographic and partial-alphabetic reading, will be observed over an extended period of time and are much more likely to be accorded

---

13 Other “immature” reading strategies, however, are commonly seen among fledgling readers of Hebrew. When novice readers have yet to master Hebrew’s complex vowel diacritics, the so-called kamats-patax kibaon phenomenon is often observed in which each (consonantal) letter is pronounced as an integral CV unit with the same default /a/ vowel (regardless of the actual vowel appearing below the consonant). For example, a CVCCVC word such as /ʀaʁɐʃ/ (butterfly) is mispronounced as /ʁaʁɐʃaʔ/—a pseudoword (see Share & Blum, 2005). A similar finding, labeled the fatha phenomenon, has also been reported in Arabic (Taouk & Coltheart, 2004).
the status of a developmental stage. Such phenomena, however, appear to be far less prevalent in regular orthographies when a compensatory head start in reading is not necessary and when phonics is the teaching method.

The nature of the preschool curriculum also seems to be a significant factor in the prevalence of logographic/partial-alphabetic phenomena. The Israeli preschool curriculum actively encourages letter learning and a variety of early reading and writing activities throughout the preschool years (ages 3–6 years), whereas the Austrian preschool curriculum explicitly discourages exposure to printed words and letter learning (Wimmer & Hummer, 1990). This downward extension of the learning-to-read phase in English-speaking communities (and other highly Americanized cultures, such as Israel) creates opportunities for “immature” strategies of the logographic/partial-alphabetic variety.

The logographic/partial-alphabetic issue, like the phonemic awareness–reading relation and the differential rates of learning to decode, can all be rallied under the same banner: When the time course of code learning is prolonged well beyond the normal span for regular orthographies (either by an unusually complex orthography or by instructional factors), each of these issues assumes proportions well in excess of the norms for conventional alphabetic orthographies.

The Role of Lexical and Supralexical Information in Decoding

An opaque spelling–sound code (or one that is functionally opaque) should increase the likelihood of nonexhaustive phonological (spelling–sound) processing among inexperienced readers, as in the case of partial-alphabetic decoding, and should also magnify general decoding ambiguity for newly encountered words, especially the more irregular ones. This applies to readers at all levels of ability and will oblige readers of English to resort to lexical and, whenever possible, supralexical or extralexical information to resolve the ambiguity (e.g., Nation & Snowling, 1998, 2004; Ricketts, Nation, & Bishop, 2007; Share, 1995; Stanovich, 1980, 2000; Tunmer & Chapman, 1998, 2006). The overall (i.e., main effect of) facilitation provided by context in English word recognition and heavy reliance by less skilled (and disabled) readers on contextual information in an effort to compensate for their lack of decoding proficiency are well established findings in English-language research (see Nation & Snowling, 1998; Stanovich, 1980, 2000). Less, however, is known about the role of lexical and supralexical factors in deciphering new words, although there is evidence that these sources of information assist the young reader in decoding unfamiliar and irregular words (Adams & Huggins, 1985; Gough & Walsh, 1991; Ricketts et al., 2007; Tunmer & Chapman, 1998, 2006). Skilled young readers may be aware that the digraph ea in the novel string heard might be pronounced in several ways, either as in heart or earth, or even hear, but they will be able to resolve the ambiguity only if they are familiar with the spoken word heard. Thus, oral vocabulary, particularly knowledge of the specific spoken forms of a large corpus of words, should be a significant factor in word learning—especially for inexperienced readers and for lower frequency and low-consistency/irregular words (Ricketts et al., 2007). Once fully-specified word-specific representations are established and new words become familiar, autonomous (or “modularized”) units, the role of context should no longer be significant (Humphreys, 1985; Stanovich, 1990). This applies to all words, regular and irregular, provided that each word (or morpheme) has a distinct graphemic configuration. (Homographs, such as wind and refuse, will necessarily remain dependent on extralexical information.) These considerations also imply a significant association between spoken vocabulary (and more generally lexical–semantic information) and early word identification, although the key factor may be knowledge of phonological forms rather than semantic information per se (see McKague, Pratt, & Johnston, 2001). The contribution of lexical and supralexical information to word recognition should be greatest among novice readers for whom a large proportion of words are not orthographically well established. The contribution should also be significant for the skilled reader encountering an unfamiliar or phonologically opaque word.

The Contribution of Lexical and Supralexical Information Depends on Transparency

Several English-language studies have confirmed a significant role for oral vocabulary and syntactic skills in word recognition, particularly for inconsistent words and for readers with impoverished decoding abilities (for reviews, see Bowey, 2005; Nation, 2005; Nation & Snowling, 1998; Woollams et al., 2007). For example, Ricketts et al. (2007) found that second graders’ expressive vocabulary predicted unique variance in exception word reading (11%) but not nonword and regular word reading after controlling for age, IQ, and decoding. Consistent with these data, a subgroup of children with poor comprehension—who were matched to a control group with good comprehension on age, IQ, and decoding efficiency—exhibited oral vocabulary weaknesses and read fewer exception words correctly. Woollams et al. (2007) reviewed an extensive literature indicating a strong association between the selective semantic deficits characteristic of semantic dementia and difficulties reading aloud low-frequency (English) exception words.

In a longitudinal study of early reading acquisition in Hebrew’s regular pointed script, Shatil and Share (2003) found no significant contribution of either oral vocabulary or syntax to a composite speed/accuracy measure of Grade 1 decoding. This supports Shatil’s (1997) hypothesis of cognitive modularity in early reading in a highly regular orthography. In another study of early (pointed) Hebrew reading acquisition among biliterate versus monoliterate Russian-speaking bilinguals, Schwartz, Leikin, and Share (2005) found that the later-immigrating biliterates possessed inferior knowledge of Hebrew syntax yet superior word recognition skills in Grade 1 thanks to phonemic awareness acquired while learning to read their native (and highly transparent) L1 Russian. Mono-

14 Although counterintuitive, this even applies to “regular” words. It appears that even the simplest novel words are not immune to English spelling–sound uncertainty. Compared with readers of more regular orthographies, English first graders exhibit far greater difficulties decoding even “regular” pseudowords for which one-to-one grapheme–phoneme correspondences should suffice (Seymour et al., 2003; Spencer & Hanley, 2003). This suggests that the notion of simple one-to-one correspondences in English is a misnomer and that the notorious one-to-many relationships for vowels create ambiguity for young readers even when confronted with seemingly simple CVC (i.e., consistent) pseudowords, such as dem.
literate bilinguals were also significantly below native Hebrew-speaking monolinguals on syntax but not reading. Similarly, Geva and Siegel (2000) found that poor semantic knowledge of Hebrew among Hebrew–English bilinguals did not impede the acquisition of (pointed) Hebrew decoding skills. Nikolopoulos et al. (2006) found that grammatical knowledge did not predict early (Grade 2) reading in regular Greek.

Further support for a Transparency × Modularity interaction comes from a within-subject study of individual differences in orthographic learning among Grade 3 Hebrew readers (Share, 2007). Share (2007) compared the same pseudoword targets (each embedded in a short passage) appearing in either shallow pointed or deep unpunctuated script. The pronunciation of the fully vowelized (i.e., pointed) targets was associated only with sublexical phonology, whereas the decoding of the same targets with the vowel diacritics deleted (i.e., unpunctuated) was related to a wide range of sublexical (phonological), lexical (morphology and semantics), and suprallexical (syntax) factors. Similar outcomes were obtained for orthographic learning. Additional support for the prediction regarding the unique role of extralexical context in deep unpunctuated Hebrew reading comes from Bentin, Deutsch, and Liberman (1990) and Deutsch and Bentin (1996). They found significant syntactic deficits among a subgroup of disabled readers having selective difficulties using sentence context to disambiguate homographs in unpunctuated text. This subgroup of readers were, nonetheless, proficient decoders of pointed text.

The same fundamental differences between pointed and unpunctuated Hebrew can also be seen in studies of context effects in the reading of vowelized and unvowelized words in Arabic (Abu-Rabia, 1997, 2001; Abu-Rabia & Siegel, 1995). Context dramatically improved (i.e., doubled) the reading accuracy of unvowelized words for both normal 10th graders (Abu-Rabia, 1997) and skilled adult Arabic readers (Abu-Rabia, 2001; see also Taouk & Coltheart, 2004, Experiment 3). Abu-Rabia (1997, 2001) has maintained that this finding is attributable to the fact that every second or third word in unvowelized Arabic is homographic. In contrast, fully vowelized Arabic, like Hebrew, is extremely regular and was unaffected by context at either age.15

These findings converge on the conclusion that both lexical and extralexical factors, such as syntax and vocabulary, are superfluous when spelling–sound relations are straightforward but are indispensable when the orthography is opaque. Generalizing the functional opacity hypothesis developed earlier in regard to the cross-linguistic phonemic awareness findings leads to the prediction that among children still learning the basic units of a regular spelling–sound code, both lexical and extralexical factors should be significant contributors to word learning. Further, semantic routes in English-language models of skilled reading (such as the DRC’s lexical–semantic subroute; M. Coltheart et al., 1993, 2001) and certain implementations of the triangle model (e.g., Plaut et al., 1996; Woolams et al., 2007) may be relevant to the pronunciation of English but not to more regular orthographies (see, e.g., Ziegler et al., 2000).

The most definitive evidence for the Transparency × Modularity hypothesis described above would be data from carefully matched English/non-English comparative studies examining the role of lexical and supralexical influences on early word identification. Until such cross-linguistic studies are forthcoming, it is worth reiterating that developmental investigations directly comparing English with highly regular German consistently reveal less modular word recognition strategies in English manifest in greater real-word substitutions (lexicalizations) when reading nonwords (Frints et al., 1998; Wimmer & Goswami, 1994), stronger lexicality and frequency effects (Frints et al., 1998; Seymour et al., 2003), and greater influences of whole-word phonology (Goswami et al., 2001). These developmental differences, moreover, are matched by cross-linguistic data from skilled adult readers summarized by the orthographic depth hypothesis (Frost, 2005; Katz & Frost, 1992). This hypothesis proposes that a phonologically opaque script calls for a greater degree of lexical involvement or “top-down shaping” than a less opaque script. These findings support the notion that readers of English, but not readers of more regular orthographies, rely more on higher order lexical factors to overcome the decoding uncertainties of English spelling.

A similar Transparency × Modularity interaction also emerges from investigations into imageability—a key semantic variable. Imageability affects exception word reading among adult skilled readers of English (Cortese, Simpson, & Woolsey, 1997; Strain & Herdman, 1999; Strain, Patterson, & Seidenberg, 1995) as well as readers with poor phonological recoding skill (Strain & Herdman, 1999). In Turkish, Italian, Greek, and Kannada—all highly regular orthographies—imageability effects are either weak or nonexistent (Karanth, Mathew, & Kurien, 2004; Mazzotta, Barca, Marcolini, Stella, & Burani, 2005; Porpodas, 1999; Raman & Baluch, 2001).

Exploiting the fact that vowel opacity in consonantal abjads can be manipulated by comparing the same word with or without vowel diacritics, Baluch and Besner (2001) found an imageability effect for opaque but not transparent words in Persian.

The overall conclusion seems inescapable: The role of lexical and supralexical factors in English word recognition for novices and experts alike does not generalize to more regular alphabetic orthographies.

Instructional Anglocentrism: The Timing and Content of Reading Instruction

In English, it takes around 3 years to reach the level of decoding mastery that is normally attained in a majority of European languages by the end of Grade 1 (Hutzel et al., 2004; Seymour et al., 2003). Thus, 3 years of learning to read (Singer, 1978) is not a myth in the code-mastery sense. Such an unusually long period of time has far-reaching implications for both the timing and content of instruction.

In consistent orthographies, the curriculum clearly reflects the full capabilities that early decoding mastery permits (see, e.g., Feitelson, 1992; McEneaney, 1997) with literature studies commencing in Grade 2. Bloomfield (1933) was by no means the first educationalist to lament “The difficulty of our spelling [which] greatly delays elementary education” (p. 500). Indeed, proponents

15 It might seem that these findings from unpointed Hebrew and Arabic call into question the present claim that phenomena, such as partial-alphabetic reading and early nonmodularity of word recognition, are largely an English-language orthographic aberration. However, it is essential to note that both Hebrew and Arabic are initially acquired as pointed scripts that have near-perfect regularity. Only several years later are the vowel diacritics dropped when children are assumed to have established visual word units in memory and words are recognized “by sight.”
of spelling reform throughout the centuries have traditionally found an ally in the teaching profession (see, Scragg, 1974, chapter 6). One of the stated goals of the "whole language" movement was to reintroduce literature that is (or was) typically postponed until after mastery of the code. In accordance with the Anglophone learning-to-decode triennium, international surveys at the primary/elementary level (invariably headed by English-language researchers) are normally conducted in Grades 3 or 4. The classic International Association for the Evaluation of Educational Achievement (IEA) studies (Elley, 1994), for example, justify the choice of Grades 3 and 4 (actually age 9 years) by asserting that "The first population [age 9 years] was at a point where most children had passed the encoding/decoding phase" (Elley, 1994, p. 6).

In several parts of the English-speaking world, formal reading instruction begins at 5 years of age rather than 6 years of age (in Britain: Duncan et al., 2006; in Victoria, Australia: Share et al., 1984; in New Zealand: Thompson, Fletcher-Flinn, & Cottrell, 1999; and in a number of U.S. states: National Reading Panel, 2000) or even at 4 years of age (see Caravolas, 2005). In countries with regular orthographies, reading instruction normally commences at 6 years of age and, in some Scandinavian countries, at 7 years of age (Elley, 1994).

Early ("emergent") literacy. In the English-speaking world, there is a strong tradition of preparatory or "emergent" literacy in the years prior to formal reading instruction in school (Lancy, 1994; Strickland & Morrow, 1989; Teale & Sulzby, 1986). Some researchers suggest that this emphasis on early literacy is a product of English spelling–sound inconsistency (e.g., Bruck, Genesee, & Caravolas, 1997; Feitelson, 1988). Ziegler and Goswami (2005) have suggested that, "owing to the difficulties in teaching the inconsistent English alphabetic code, English-speaking countries have begun to teach reading and ‘pre-reading’ skills at younger and younger ages" (p. 14). Bruck et al. (1997) have maintained the following:

Preliteracy skills are deemphasized in language communities that learn a relatively transparent orthography because it is assumed that reading acquisition is relatively straightforward. Conversely, English-speaking communities may place a great deal of emphasis on preliteracy skills because it is difficult for children to learn English, one of the most inconsistent and less transparent alphabetic languages. (p. 160)

In most countries with regular orthographies, there is little or no reading preparation before formal schooling (see, e.g., de Jong & van der Leij, 1999; Elbro, 2006; N. C. Ellis et al., 2004; Landerl, 2000; Mann & Wimmer, 2002; Thorstad, 1991; Wimmer, Mayringer, & Raberger, 1999). According to Wimmer, Landerl, and Frith (1999), “[I]n [German] kindergarten[s] there is no reading preparation at all” (p. 36). This sweeping assertion has empirical backing; Mann and Wimmer (2002) found that English preschoolers were far superior in naming letters, word reading, phonological awareness, and knowledge of environmental print compared with matched German speakers. This head start, however, was rapidly eclipsed by the German speakers by the end of Grade 1 (bearing out Bruck et al.’s, 1997, observations cited above). The developmental trajectory consisting of an early head start by English learners that is soon surpassed (in this case by Turkish readers) has been reported by Durgunoglu and Oney (1999) for letter knowledge. Bruck et al. (1997) also found superior letter name knowl-

edge among English-speaking Canadians compared with French-speaking Canadians. Moreover, the researchers reported more story reading among English parents despite higher educational levels among the French-speaking mothers. It is worth noting that many educationalists discussing beginning reading instruction in consistent scripts (e.g., Feitelson, 1988) have eschewed any form of reading instruction, formal or informal, prior to first grade.

Whole language and the English spelling–sound code. Snow and Juel (2005) have maintained that much of the debate on the pedagogy of reading has revolved around the “villain” of orthographic depth (p. 505). The vagaries of the English spelling–sound code have assumed a prominent place in the antiphonics rhetoric of the whole language movement and perhaps even more so in its popularization.16

According to F. Smith (1978), one of whole language’s founding fathers, “the ‘rules’ relating spelling to the sounds of speech are inordinately complicated and unreliable in English . . . the effort to read through decoding is not only futile but unnecessary” (p. 2). Becoming a fluent reader, F. Smith asserts, means learning to become less reliant on visual (i.e., print) information and more reliant on nonprint nonvisual information (i.e., prior knowledge and the generation and testing of predictions on the basis of prior knowledge). “The spelling–sound correspondences of English are so confusing that in my judgment children who believe they can read unfamiliar words just by ‘blending’ or ‘sounding’ them out are likely to develop into disabled readers” (F. Smith, 1978, p. 55).

Goodman, although somewhat less truculent than F. Smith about the problems of English spelling, has claimed that the use of spelling–sound relationships is the least helpful of the three cue systems (grapho-phonetic, syntactic, and semantic). “Readers are able to use syntactic and semantic cues to such a considerable extent that they need only minimal graphic cues in many cases” (Goodman, 1986, p. 64). “Developing readers must be encouraged to predict and guess as they try to make sense of print” (Goodman, 1986, p. 39).

It is not difficult to show that F. Smith (and, to a lesser extent, Goodman) exaggerates English spelling irregularity by ignoring morphemic, positional, and contextual regularities (see Carney, 1994); however, similar concerns have been voiced by many authors, not only those identified with the whole language approach. Bruck et al. (1997), for example, suggested that “for English, a strict phonics approach may be less likely because of the irregularities and inconsistencies in the writing system” (p. 149). In more regular orthographies, beginning reading instruction is more phonics oriented than in English (see, e.g., Bruck et al., 1997; Caravolas & Bruck, 1993; Duncan et al., 2006; Haghtvet & Lyster, 2003; Hutzler et al., 2004; Landerl, 2000; Lundberg, 1994; Mann 16 I do not wish to imply that whole language was solely the outcome of despair about English spelling. This was only one among a number of factors, including the ascendancy of functionalist/constructivist perspectives on language and learning in general, as embodied in the work of theorists such as Halliday, Piaget, and Vygotsky (see Goodman, 1989; Goodman & Goodman, 1979), teacher empowerment issues (Goodman, 1989), literature-based instruction (the latter itself partly a response to the problem of the protracted learning-to-read phase in English), an early emphasis on writing, motivation and affect, and the specific (now-refuted) claim that inefficient readers overattend to visual information (Goodman & Goodman, 1979).
English word reading instruction often involves a combination of code instruction (phonics) and “sight-word” instruction employing “look-and-say” or whole-word methods in regular classrooms (N. C. Ellis et al., 2004; Foorman et al., 2004; Mann & Wimmer, 2002; McEneaney, 1997; Nikolopoulos et al., 2006; Reitsma & Verhoeven, 1990).17 English orthography is exceptional in yet another regard. With the possible exception of the apostrophe and the occasional loan word (e.g., naïve), diacritics are foreign to standard printed English. However, diacritical markings, of the kind familiar to English readers only in a pronouncing dictionary, are a common feature of the orthographic landscape of all major language families (Daniels & Bright, 1996). Throughout the history of writing, diacritics have proven a popular device for enabling an orthography to represent additional phonemes without adding new graphemes.

Diacritics represent a kaleidoscope of orthographic embellishments employed to modify and supplement basic graphemic information, such as marking tones (e.g., in pinyin in China; Mair, 1996); indicating voicing (e.g., in Japanese kana; J. S. Smith, 1996, p. 212); marking stress (e.g., in Greek; Threatte, 1996); and signaling vowel quality (e.g., in Arabic; Abu-Rabia & Taha, 2006), vowel duration (e.g., in Hungarian; Csépe, 2006), vowel nullification (e.g., in Tamil; Aaron & Joshi, 2006), nonnative phonemes (e.g., in Kannada; Prakash & Joshi, 1995), and much more. Among European tongues, the Slavic scripts are readily distinguished by an efflorescence of diacritics. These mark a host of phonetic distinctions for both vowels and consonants, the ubiquitous haček being the best known. In pointed (fully vowelized) Semitic scripts, the number of diacritics often exceeds the number of letters (Share & Levin, 1999) and indicates a multitude of vowel and consonant alternations (Ravid, 2006).

As a case in point regarding the dominance of the English-language research agenda, the ubiquity of Hebrew vowel diacritics has provided Hebrew-language researchers with an invaluable tool for examining the role of phonological information in word reading by permitting within-item comparisons (the same word with and without vowel diacritics) that are not possible in English. This unique feature of Semitic orthography has been exploited almost exclusively to address one of the central issues on the English-language research agenda: the role of phonology in word recognition. As such, almost all of this work (see, e.g., Shimron, 1993) has regarded diacritics as a means to an end, not as the central focus of investigation, until very recently (Gur & Share, 2007; Ravid, 2006).

Diacritical markings are also a popular tool for enhancing the decipherability of nontransparent scripts for young readers. Simplified orthographies used as stepping stones for the learner (“pédagogies”) often make extensive use of diacritics (e.g., i.t.a.) and are a standard feature of consonantal abjads, such as Arabic, Hebrew, and Persian. As noted earlier, these feature both vowelized and unvowelized versions. In these dual-purpose scripts, vowelization by means of diacritics is the norm during the early years of schooling (Abu-Rabia & Taha, 2006; Baluch & Besner, 1991; Ravid, 2006; Share & Levin, 1999) after which the diacritics are discarded. With respect to the contemporary English-language research agenda, the topic of diacritics is, unsurprisingly, a nonissue.

Conclusions

Commonalities and Universals

The current review has focused on disparities between English and other (primarily alphabetic) orthographies, but commonalities must also be acknowledged. The transition from unfamiliar to familiar in the case of words and from novice to expert in the case of readers offers a useful framework for conceptualizing several universal aspects of reading.18 Every word is visually unfamiliar at some point; all orthographies must therefore supply readers with a means for independently deciphering new words that also lays foundations for future rapid memory-retrieval processes. This leads to the first linguistic universal or, more precisely, phonological universal of writing systems (DeFrancis, 1989; Hanley et al., 1999; Ho & Bryant, 1997; Perfetti, 2003; Perfetti, Liu, & Tan, 2005; Perfetti, Zhang, & Berent, 1992; Pollatsek et al., 2000). Orthographies must directly represent a finite set of recombiant

17 Yet another by-product of English’s outlier orthography may well be the huge instructional diversity that characterizes English teaching methods (see, e.g., Aukerman, 1984). By contrast, in the Netherlands, the same structured phonics program can be found in 8 out of 10 first-grade classrooms (Reitsma & Verhoeven, 1990). Substantial environmental variability also has consequences for estimates of gene/environment contributions to reading across consistent versus inconsistent orthographies. Bishop (2001) cited two British twin studies indicating that shared environment accounts for most of the variability among (English-speaking) dyslexics. She hypothesized that variation in instruction may be a key factor here.

18 It should be emphasized that the (old/new) unfamiliar-to-familiar/novice-to-expert dualism proposed here as an alternative to the dominant Anglophone regular–irregular dualism does not presume to compete with the sophistication and level of detail of current computational models but rather to suggest a different conceptual starting point for the model-building enterprise. As such, the unfamiliar-to-familiar/novice-to-expert framework presents only general and necessarily underspecified guidelines given that this new point of entry to the model-building process is quite a distance “upstream” from the much more highly developed or “downstream” efforts of contemporary computational modelers. (Ans et al., 1998, multitrace computational model of reading might be one specific instantiation of a general class of models that embodies (at least in part) this alternative dualism as distinct from dual-route models [e.g., DRC and connectionist dual-process] that embody a fundamentally different set of axioms.)
sublexical units of speech—the potentially infinite set of word meanings can be represented only indirectly (A. M. Liberman, 1992; Mattingly, 1985). An appreciation of and facility for manipulating those same units, phonological awareness, is therefore a universal ingredient in successful reading acquisition in all languages and scripts (Caravolas, 2005; Cheung, McBride-Chang, & Chow, 2006; Cosset al., 1988; Duruguoglu & Oney, 1999; Geva et al., 1993; Hanley, 2005; Landerl et al., 1997; Share & Levin, 1999). Inefficient processing of print-to-sound connections, expressed as inaccuracy or dysfluency (or both), is another universal feature of reading difficulties (Caravolas, 2005; Ho & Lai, 1999; Kim & Davis, 2006; Landerl et al., 1997; Paulesu et al., 2001; Rack et al., 1992; Vellutino et al., 2004; Ziegler, 2006).

The second element of the unfamiliar-to-familiar/novice-to-expert dualism speaks to the essence of reading skill. In accordance with the orthographic maxim of morphemic distinctiveness and the self-teaching inherent in the process of deciphering unfamiliar words (Share, 1995, 2008), the advancing reader eventually assembles a large stock of instantly familiar words, each recognized as an integrated autonomous unit. The item-by-item accumulation of well-unitized orthographic entries (localized or distributed) ultimately requires years of daily print exposure. This common end-state of reading “fluency” looks remarkably similar across diverse languages and orthographies. These commonalities notwithstanding, the discommonalities—as I, and others have attempted to argue (Frost, 2005; Katz & Frost, 1992; Seymour, 1999; Kim & Davis, 2006; Landerl et al., 1997; Paulesu et al., 2001; Rack et al., 1992; Vellutino et al., 2004; Ziegler, 2006).

The history of psychology, like the history of science, has witnessed a colorfull pageant of ambitious theories (e.g., Freud, Piaget, Skinner, Newell & Simon), each aspiring to the status of grand unifying theories of human behavior. Over the course of time and painstaking empirical research, each theory has been whittled down to modest proportions but still offers valuable insights into a limited range of domain-specific phenomena. All theories are necessarily constrained by a finite set of observations bound by time and circumstance; thus, a complete science of reading ultimately requires a deep understanding of the universal and script-specific nature of reading and writing across all writing systems. Reading science cannot be founded on a single, outlier orthography. The sheer volume of English-language reading research and the prestige that English currently enjoys in contemporary science, technology, economics, and culture have meant that the Anglophone research agenda often dictates the research questions pursued by researchers the world over. We persist in comparing other orthographies with English (see, e.g., August & Shanahan, 2006). Findings from non-English orthographies are invariably evaluated in the light of Anglophone theory and data. Non-English research is too often motivated by the question of whether an English-language theory or finding extends to language X or to orthography X, where X is regarded as a special case. I contend that it is English that warrants the special case status and that the outcomes from non-Anglophone studies are likely to offer a better approximation to the global norm. At the very least, the wholesale transplantation of conclusions from reading research in English is ill-advised. Reading science can afford to be neither Anglocentric nor Eurocentric. To see the larger picture of how reading works (Perfetti, Liu, & Tan, 2005, p. 57) calls for comparative analyses. It cannot be informed by a sole highly atypical exemplar. This said, however, a science of reading must account for reading behavior and development in both conventional and exceptional orthographies to be accorded the status of universal.

Postscript. The unrivalled status that English enjoys today as the global language (Crystal, 2003) has brought a proliferation of English borrowings and associated spellings (often nonnaturalized). Ironically, many highly regular (Roman-based) scripts (perhaps hundreds) are now finding significant numbers of borrowed English spellings (exceptions, no exception) in their orthographic lexicon. The wheel may be coming full circle, with the extraordinarily hospitable English lexicon now dispensing orthographic largesse in every corner of the globe. Could this one outlier orthography be shifting the global norm of spelling–sound correspondence among Roman-based scripts? Whether English spelling exceptionality is on the wane, no grand unifying theory in the reading field can afford to overlook such a ubiquitous outlier.

This appears to apply even to nonalphabetic Chinese characters. Given that most Chinese words are polysyllabic, the finite number of mono-syllabic phonetics (approximately 240 according to Taylor & Taylor, 1995, or 1,000 if tones are included) might be construed as a recombinant (quasi-)“sublexical” unit. There are also several lines of evidence suggesting that even these phonetic elements are not indivisible whole units and that learning to read Chinese is not merely a process of rote learning of unsegmented whole-character-to-whole-syllable associations (Paradis, 1989; Wang, 1973). First, a number of studies have demonstrated that Chinese children have phonological awareness, and although the exact nature of the unit is yet to be resolved (see Hanley, 2005), the available evidence is in agreement as regards to the sublexical, if not phonemic, nature of this awareness (Hanley, 2005; Ho & Bryant, 1997; Siok & Fletcher, 2001; So & Siegel, 1997). Second, there are reliable data showing regularity effects in Chinese (Perfetti, 2003; Seidenberg, 1985; Shu et al., 2000). In a related finding, R. C. Anderson et al. (2003) found second- and fourth-grade Chinese readers were able to make use of partial (i.e., sub-syllabic) information in the phonetic to learn and remember the pronunciation of novel semantic–phonetic compound characters (see also discussion of the ancient fangie principle; Leong, 1995). Among these novel characters were phonetics (termed onset-different) that shared only the “final” (rime) unit. Furthermore, performance on these semiphonetic elements was superior to characters with no common phonetic elements. Third, errors also confirm the influence of sublexical (initial/final) information. It should be acknowledged, however, that unlike alphabetic orthographies, Chinese phonetics are only one piece of a two-part puzzle; meaning-based radicals have also been shown to be an important source of information in learning new characters (Shu & Anderson, 1997). Perhaps the strongest evidence for the universality of the phonological principle is the introduction of alphabetic scripts, such as pinyin, designed to assist children in learning new characters—a self-teaching device par excellence.

References


entropy matters—Letter-to-phoneme mappings in seven languages. Reading and Writing, 18, 211–229.
Coltheart, V., Patterson, K., & Leahy, J. (1994). When a ROWS is a ROSE: Phonological effects in written word comprehension. Quarterly Journal
of Experimental Psychology: Human Experimental Psychology, 47(A), 917–955.
Foley, B. E., & Pollatsek, A. (1999). Phonological processing and reading abilities in adolescents and adults with severe congenital speech impair-


and without hearing impairments: Some insights from an associative learning task. *Reading and Writing*, 17, 823–845.


Received September 4, 2007
Revision received February 13, 2008
Accepted February 25, 2008