12 Phonological technologies: reading and writing

Core knowledge systems outline not only our early, instinctive, and universal understanding of the world but also provide scaffolds for subsequent learning. Like the core systems of number, physics, and social knowledge, our instinctive phonological knowledge sets the stage for the cultural invention of reading and writing. This chapter outlines the intimate link between early phonological competence and those later "phonological technologies". We will see that all writing systems – both conventional orthographies and the ones invented spontaneously by children – are based on phonological principles. Reading, in turn, entails the automatic decoding of phonological structure from print. Skilled reading recruits the phonological brain network that mediates spoken language processing. Moreover, dyslexia is typically marked by hereditary deficits to phonological processing and phonological awareness. The role of instinctive phonology as scaffolds for reading and writing is in line with its view as a system of core knowledge.

12.1 Core knowledge as a scaffold for mature knowledge systems

In previous chapters, we have seen that phonological systems manifest a unique, potentially universal design that is evident already in early development. The special design of the phonological system is in line with the characteristics of core knowledge systems documented in numerous other domains, including knowledge of number, agency, space, and morality (Bloom, 2010; Carey, 2009; Carey & Spelke, 1996; Hamlin, Wynn & Bloom, 2010; Hamlin et al., 2007; Hauser & Spelke, 2004; Spelke, 2000). These early knowledge systems each include distinct representational primitives and combinatorial principles that are innate, universal and domain specific. For example, infants as young as 4 months of age manifest rudimentary knowledge of number – they can represent the precise number of up to four objects (larger numbers are encoded approximately),

PROOF

C:/ITOOLS/WMS/CUP-NEW/3328241/WORKINGFOLDER/BERE/9780521769402C12.3D 281 [280–306] 28.8.2012 6:16AM

and they can perform addition and subtraction operations on such small sets. In the domain of physics, young infants possess intuitive knowledge that leads them to expect objects to move cohesively (without disintegrating) and continuously (without jumping from one point to another and without intersecting other objects) as a result of contact with other objects. Other principles of morality might underlie 3-month-old infants' preference for social "helpers" (a character helping another climb up the hill) to "hinderers" (a character who interferes with the climber's efforts).

While these early, intuitive knowledge systems continue to play a role throughout development, as the child matures they gradually give rise to new bodies of knowledge that differ from their predecessors in their contents and expressive power (Carey, 2009). For example, the core number system available to infants is limited in size – it can implicitly encode precise numerosity of sets of up to four objects. Adults, in contrast, can compute the numerosity of any set by relying on a later-emerging system of recursive number that develops on the heels of the primitive number systems available to infants and animals. In a similar manner, the child's early concepts of object and motion eventually give rise to elaborate scientific theories of physics (Spelke, 1994), and infants' basic intuitive moral sense lays the foundation for moral systems that apply generally, to both kin and stranger (Bloom, 2010).

Unlike their intuitive innate predecessors, those later theories and inventions are by no means instinctive or universal, as different cultures vary in their scientific and technological advance as well as their moral codes. While the early core number systems, for example, are present universally, in any infant, the later system of recursive number depends on specific linguistic experience with number words and quantifiers, and consequently, people deprived of such linguistic devices – either because their language lacks them (Gordon, 2004) or because they lack access to a language (Spaepen et al., 2011) – do not develop the recursive number system. Likewise, scientific discoveries and moral theories are the product of deliberate reasoning and the intense research of a select few individuals, rather than the outcome of biological maturation available universally. But although the elaborate cultural discoveries, theories, and technologies of adult communities clearly differ from the intuitive universal and innately based early systems of core knowledge, these two kinds of knowledge are nonetheless linked inasmuch as several of those later discoveries develop on the heels of their ontogenetic predecessors.

Just as our intuitive core knowledge of number and object gives rise to mature scientific theories of mathematics and physics, so does the core system of phonology form the scaffold for a cultural technological invention – the invention of reading and writing. Indeed, reading and writing are intimately linked to phonological knowledge (Liberman, 1973; Perfetti, 1985). As I next demonstrate, all fully developed writing systems encode concepts using phonological

C:/ITOOLS/WMS/CUP-NEW/3328241/WORKINGFOLDER/BERE/9780521769402C12.3D 282 [280–306] 28.8.2012 6:16AM

means. The precise link between the writing system and phonological patterns varies – some writing systems encode phonemes, whereas others represent syllables. But the reliance on a phonological system is common to them all. And since writing systems are only part inventions and mostly discoveries – the discovery of spoken phonological patterns – the designs of writing systems tend to converge across cultures. In fact, such phonologically based designs reemerge spontaneously in the rudimentary writing systems that are routinely invented by many children.

Not only does core phonology form the basis for writing, but it also constrains reading. Like many skilled readers, you, the reader of this book, might decode the words printed on this page automatically and effortlessly, with no awareness of their phonological structure. For this reason, most people believe that they identify printed words in much of the same way they identify traffic signs – by directly mapping visual symbols onto concepts. But a large experimental literature shows that this popular belief is in fact mistaken. All readers, both beginners and skilled, routinely go through the extra step of mapping graphemes onto phonological representations. As in the case of writing, the phonological representations decoded in reading vary in grain size depending on multiple factors (e.g., the writing system, the familiarity with the word, and the specific experimental conditions). Nonetheless, some level of phonological decoding is always present (Perfetti et al., 1992), and it is demonstrably shaped by the same phonological principles documented in spoken language. Skilled reading is thus strictly constrained by phonological competence. Conversely, when people are unable to effectively encode the phonological structures of spoken language, dyslexia typically ensues.

Why would a cultural invention such as reading pay its debt to phonology? Unlike math and physics, reading and writing are just invented codes, not theories of the physical world. While an arbitrary theory of physics cannot be maintained in the face of conflicting empirical evidence, and physically improbable technologies are bound to go extinct, phonologically arbitrary writing systems are amply feasible, and direct "visual" encoding of such systems is certainly conceivable. In fact, on some accounts, it is even likely. But precisely because phonologically arbitrary writing and reading is logically possible, the fact that such systems are disfavored suggests internal constraints on their design. The hypothesis of core phonological knowledge accounts for these facts. Like its sister systems of core knowledge – number, physics, and morality – reading and writing might be grounded in core knowledge. The link between the instinctive, early phonological system and reading and writing, in turn, provides converging evidence for the existence of a core system in phonology. This chapter reviews the evidence linking reading and writing to core phonology. These observations demonstrate that writing systems – both conventional and invented ones – are all based on phonological principles, that their decoding in reading recovers phonological structure from print, and that deficits to core phonology are linked to reading disability.

12.2 Writing systems recapitulate core phonology

12.2.1 Conventional writing systems

Writing systems are inventions that have independently emerged at least three times in the history of the humanity. The first writing system was devised by the Sumerians roughly 5,000 years ago, about 1,500 years later, a second system was invented by the Chinese, and a third system was independently devised about 2,000 years ago by the Mayans (Rogers, 2005). While most inventions are judged for their originality, in the case of writing, the similarity among those independent inventions is even more striking than their differences. Each of these ancient systems includes some method of phonological organization, and the reliance on phonological encoding has since been preserved in practically every fully developed writing system (for one possible exception, Bliss, see Rogers, 2005).

Chinese characters, for example, encode syllable-size phonological units. Indeed, the number of characters expressing a single Chinese word depends on the number of syllables: Monosyllabic words are conveyed by a single character and disyllabic words by two, and this link holds irrespective of whether any given syllable is mono-morphemic or bi-morphemic (Rogers, 2005; see 1). Thus, the disyllabic words *shān hú* ('coral') and *tiě lù* ('railway') are each expressed by two characters, even though the former is mono-morphemic whereas the latter is morphemically complex.

(1) Chinese characters correspond to syllable-size units (not morphemes; examples from Rogers, 2005; *Chinese Character Dictionary*, 2010)
   a. Monosyllabic words are expressed by a single character:
      wǒ ('I') 我
      hǎo 好 ('good')
   b. Disyllabic words are expressed by two characters:
      (i) monomorphemic
         hú dié ('butterfly') 蝴蝶
         shān hú ('coral') 珊瑚
      (ii) di-morphemic
         tiě lù ('railway'; = *tiě* 'iron' + lù 'road') 鐵路
         zì diǎn ('dictionary'; = *zì* 'character' + *diǎn* 'standard') 字典

Not only does each Chinese character correspond to a single syllable, but segments that form a single syllable may be encoded by the same character even when meaning varies (see 2). For example, the words for 'horse' and 'mother' both share the same syllable, *ma*, encoded by a common orthograhpic character

C:/ITOOLS/WMS/CUP-NEW/3328241/WORKINGFOLDER/BERE/9780521769402C12.3D 284 [280–306] 28.8.2012 6:16AM

(馬). This is not to say that the phonological representation of Chinese characters is fully predictable from print. Although some characters (about 25 percent of all characters) convey phonological information (segmental and tone) fully and reliably, others carry only partial phonological information, and some characters (estimated at 33 percent of all characters) carry no useful phonological at all (DeFrancis, 1989: 113). But the fact that such links nonetheless exist demonstrates that phonological principles play a role in the design of this orthography.

(2) Segments that form a single syllable are encoded by the same character (examples from DeFrancis, 1989; Rogers, 2005; *Chinese Character Dictionary*, 2010)

a. /ma/ homophones:
  媽 'mother' ma1
  馬 'horse' ma3

b. /jiao/ homophones:
  僥 *jiǎo* 'lucky'
  澆 *jiāo* 'to water'

While Chinese encodes syllable-size characters, other writing systems contrast among finer-grained phonological units – moras, segments, and even feature-size units. Moras are units of prosodic weight, and weight, in turn, typically depends on the structure of the rhyme: CV syllable counts for a single mora, whereas CVV and CVC units count for two. In the Japanese Kana systems (the 'plain' Hiragana and 'side Kana' Katakana), most characters correspond to monomoraic syllables. Accordingly, Katakana indicates monomoraic, CV syllables by a single symbol whereas bimoraic CVV and CVN (N=nasal) syllables are indicated by two symbols. For example, the initial monomoraic syllable in Toyota (see 3) is transcribed by a single Hiragana character, と, whereas the bimoraic word /too/, 'ten' comprises two symbols とお – the initial と symbol from Toyota, followed by an additional symbol for the second mora (see 3).

(3) The expression of moraic contrasts in Hiargana
  とよた Toyota /toyota/ <to.yo.ta>
  とお ten /too/ <to.o>

Moving down the inventory of phonological primitives to the level of the phonemes, we arrive at the familiar alphabetic writing systems – systems that use graphemes to encode phonemes. Some alphabets, such as English, encode both consonants and vowels, whereas consonantal orthographies such as Hebrew encode mostly consonants to the exclusion of most vowels. The examples in (4) are all morphologically relatives, derived form the root /ktv/, 'writing,' a fact depicted in the orthography by their common consonant letters (כתב).

(4) The Hebrew consonantal orthography
  כתב /katav/ 'he wrote'
  כתב /ktav/ 'handwriting'
  תבכ /ktiv/ 'spelling'

C:/ITOOLS/WMS/CUP-NEW/3328241/WORKINGFOLDER/BERE/9780521769402C12.3D 285 [280–306] 28.8.2012 6:16AM

Finally, several orthographies use symbols to encode sub-phonemic feature distinctions. Korean phonology, for example, contrasts obstruents in terms of their aspiration and the tense-lax dimension. The Hangul orthography, in turn, expresses these distinctions: aspiration is marked by adding a stoke to unaspirated counterparts, and tenseness is expressed by reduplicating the corresponding lax consonant (Rogers, 2005; see 5). Moreover, Hangul conveys the syllabic constituency of the phonemes in terms of their systematic spatial arrangement: onsets are obligatorily encoded either by a consonant, or, in the case of vowel-initial syllables, by a dummy character ㅇ, vowels (including on-glides, *y* and *w*) are encoded either to the right of the onset or below it (depending on the direction of their main stroke – vertical or horizontal), and coda consonants are indicated at the bottom of the cluster (see 6).

(5) The representation of aspiration and tense-lax contrasts in Hangul
 t ㄷ, t^h ㅌ, tt ㄸ
 p ㅂ p^h ㅍ ppㅃ
 k ㄱ, k^h ㅋ, kk ㄲ

(6) The spatial depiction of syllable structure in Hangul (from Simpson & Kang, 2004) 교실 /kyo.sil/ 'classroom'
 교 /kyo/
  onset: k ㄱ
  nucleus: yo ㅛ
 실/sil/
  onset: s ㅅ
  nucleus: i ㅣ
  coda: lㄹ

These examples make it plain that conventional writing systems are based on phonological principles – they encode the same set of phonological primitives attested universally in phonological systems – syllables, moras, segments, and features – and, in some cases, writing even expresses their structural roles in the syllable.

12.2.2 Invented spelling systems

The link between writing system and phonology is not merely a diachronic fact about the evolution of writing systems. Rather, it is a vital synchronic phenomenon that reemerges time and time again in the spellings invented by children. Children tend to spontaneously invent spelling systems of their own based on rudimentary familiarity with the letters of their adult community (Chomsky, 1976; Read, 1971; Treiman & Cassar, 1997). Being invented systems, these spellings are bound to differ from conventional systems. But precisely because such "misspellings" diverge from the adult model, they provide a window into

C:/ITOOLS/WMS/CUP-NEW/3328241/WORKINGFOLDER/BERE/9780521769402C12.3D 286 [280–306] 28.8.2012 6:16AM

the productive knowledge that shapes the formation of those systems. As it turns out, the relevant knowledge is phonological.

The role of phonological knowledge in shaping the child's invented spelling results in two broad contrasts compared to the adult's system: under-specification – cases in which the child omits some of the phonological distinctions present in the adult's system, and over-specification – cases in which children specify some phonological contrasts that adults fail to express in their mature spelling system.

Under-specification takes multiple forms. Children's early spellings of English, for example, often conflate the contrast between tense and lax vowels: Children first use tense vowels (whose sounds are familiar to them from the letters' names) to express their corresponding lax counterparts (e.g., they use *E* to spell both /i/ and /ɪ/), and once they learn the proper way to express lax vowels, they subsequently over-generalize those spellings to express tense vowels (they use ***i*** to spell /i/). Another common error is the omission of vowels before sonorant segments (e.g., *tiger*➔ TIGR).

(7) Phonological under-specification in invented spelling (from Read, 1971)
  a. Failure to contrast tense and lax vowels:
    FEL (feel)
    FES (fish)
    SIKE (seek)
  b. Failures to specify the vowel:
    TIGR (tiger)
    DIKTR (doctor)

But these two types of errors are neither careless nor arbitrary: Both patterns reflect productive phonological knowledge. Indeed, children do not randomly confuse any two-vowel phonemes, but they specifically conflate vowels that are matched for height and differ only on their tenseness. In so doing, they demonstrate that they know that these phoneme pairs share a feature. Moreover, the conflation of the tense-lax dimension in the child's spelling mirrors a phonological abstraction present in conventional English orthography (e.g., *extreme–extremity*; *divine–divinity*), but this convergence most likely emerges independently – it is unlikely that children simply imitate the conventional adult spelling, as their invented spellings precede spelling instruction. Similarly, children's tendency to omit vowels before syllabic consonants (e.g., TIGR, DOCTR) are systematic – children are reliably more likely to omit such vowels compared to ones preceding nonsonorant consonants (e.g., *sal**a**d*, *bask**e**t*; Treiman, 2004). Although such spellings happen to counter the obligatory encoding of vowels in all spelled syllables, they are perfectly consistent with the phonology of English, where sonorant consonants can form a syllable of their own (e.g., the *r* in *tiger*, pronounced [tajg]).

C:/ITOOLS/WMS/CUP-NEW/3328241/WORKINGFOLDER/BERE/9780521769402C12.3D 287 [280–306] 28.8.2012 6:16AM

(8) Phonological over-specification in invented spellings (from Read, 1971)
a. Regular suffix:
MARED (married)
HALPT (helped)
b. Affrication:
CHRIE (try)
JRGAIN (dragon)
c. Flaps:
LADAR (letter)
PREDE (pretty)

The systematic phonological basis of children's "misspellings" is also evident in cases where their spellings specify phonological distinctions that are attested in English phonology but are unspecified in conventional adult spelling system (i.e., over-specification). English spelling conflates the voicing distinction between the suffix in *married* and its voiceless counterpart in *walked*, but children obey this phonological contrast in their invented spellings. Similarly, by expressing the intervocalic *t* (e.g., in *letter*) by a D, children approximate its realization as a flap, rather than a [t]. In all these cases, children's spellings reflect accurate renditions of the phonology that are absent in the adult's systems. Although these examples are limited inasmuch as they are all confined to a single language, English (Share, 2008), there is some evidence that the spontaneous extraction of phonological organizational principles also extends to nonalphabetic writing systems (Nag et al., 2010).

Summarizing, then, all full writing systems deploy some phonological organizational principles. The intimate link between phonology and spelling is present in conventional orthographies that evolved from three independent phonological writing traditions, and it recapitulates in the spelling systems that are routinely and spontaneously invented by young children. Although all writing systems are ultimately inventions that utilize visual symbols, at their core, they are discoveries – the discovery of one's phonological system. Thus, writing systems are systems of visible speech (DeFrancis, 1989).

12.3 Reading recovers phonological form from print

That writing recapitulates phonology is intriguing, but perhaps not entirely surprising. After all, writing systems encode language, and all human languages are known to exhibit phonological patterns that routinely generalize to novel forms. If writing systems are to keep up with the vast richness of spoken phonological forms and their constant expansion by innovations and borrowings, then some productive phonological principles must be incorporated in the writing system itself. Remarkably, however, the phonological reflexes of written language are

PROOF

C:/ITOOLS/WMS/CUP-NEW/3328241/WORKINGFOLDER/BERE/9780521769402C12.3D 288 [280–306] 28.8.2012 6:16AM

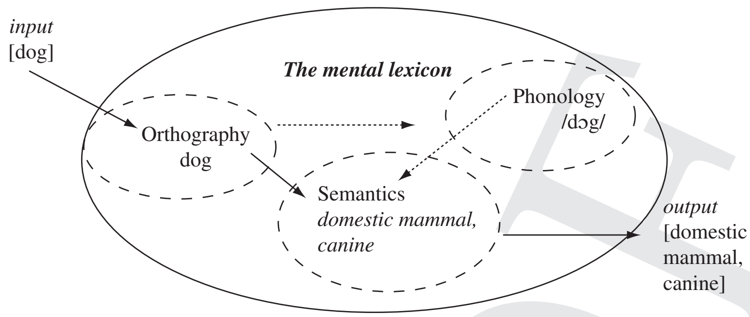

Figure 12.1 Lexical access from print

evident not only at its encoding, in writing, but even during its online decoding – in reading.

Unlike the phonological encoding of print, the reliance of reading on phonology is hardly expected. While one must rely on phonological principles to decode novel words – the *ipads*, *faxes*. and the countless other gadgets that are constantly added to our lives – the grand majority of words are highly familiar. For such words, readers have securely stored in their memory representations that link the word's spelling with its meaning. To recognize a familiar printed word, that is, to associate it with some word's meaning, stored in the mental lexicon – readers could simply retrieve the meaning directly from its spelling (see the heavy continuous arrows in Figure 12.1). Decoding the word *dog* would essentially proceed along the same lines as any other nonlinguistic sign – traffic signs, faces, and scenes.

The possibility of non-phonological decoding of familiar English words is difficult to grasp precisely because we automatically retrieve the phonological form of words from print. But a brief example from a foreign orthography reminds us that phonological decoding is by no means necessary. English readers unfamiliar with the Hebrew orthography can easily learn the meaning of *ëìá* by associating it with an image (see Figure 12.2), oblivious to the fact that its phonological form happens to be /kelev/. In a similar manner, readers could decode most printed text by associating visual symbols directly with their meaning. But countless studies have shown that this is not what readers typically do. Rather than directly mapping graphemes and meanings, readers of all languages routinely rely on some method of phonological decoding. The precise method varies across orthographies, but the reliance on some form of phonological decoding appears to be universal. The conclusion that reading entails obligatory phonological processing underscores the intimate link between reading and the core system of phonology. In what follows, I briefly review some of the evidence for phonological decoding at both the sentence and single-word levels.

C:/ITOOLS/WMS/CUP-NEW/3328241/WORKINGFOLDER/BERE/9780521769402C12.3D 289 [280–306] 28.8.2012 6:16AM

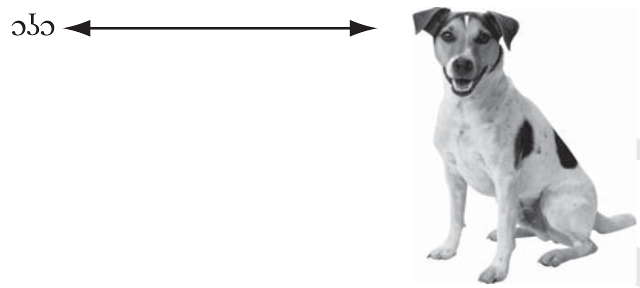

Figure 12.2 Reading without phonology

12.3.1 Phonological decoding in the silent reading of sentences

For skilled readers, the silent decoding of printed sentences typically appears just that – silent. But appearances can be misleading, and tongue-twister sentences make this fact patently clear. Tongue twisters are painfully hard to utter aloud. But remarkably, this challenge persists even when reading is silent. Numerous studies have shown that tongue-twister sentences (see 9) are harder – they take longer to (silently) read and they are subsequently harder to recall compared to control sentences matched for syntactic and semantic structure (e.g., Keller et al., 2003; Kennison et al., 2003; McCutchen & Perfetti, 1982; McCutchen et al., 1991; Zhang & Perfetti, 1993.)

(9) Tongue-twister sentences and controls:

A tongue-twister sentence:	the taxis delivered the tourist directly to the tavern
Control:	the cabs hauled the visitor straight to the restaurant

Several observations suggest that the difficulties with tongue twisters are not visual confusions due to letter repetitions. First, the same difficulties obtain even when the repeated phonological elements (phonemes, features) are expressed by different letters (e.g., *t* vs. *d*, in English; McCutchen & Perfetti, 1982). Second, the interference from tongue twisters interacts with a secondary phonological task of digit recall (McCutchen et al., 1991). In the experiment, people are first asked to memorize a set of five numbers (e.g., 2, 12, 25, 22, 29), they are next presented with a sentence – either tongue twister or control which they are asked to judge for meaning – and they are finally asked to recall the set of digits in the order they were presented. Results show that the tongue-twister effect depends on the phonological similarity between the sentence and the digits. For example, tongue twisters repeating an alveolar stop (*the taxis delivered the tourist* . . .) are harder to read in the context of phonologically similar digits (e.g., the initial alveolar stop in 2, 12, 25, 22) compared to less similar controls (e.g., the initial alveolar fricatives in 6, 7, 66), and the similarity between the sentences and numbers impairs number recall as well. A third piece of evidence demonstrating that the tongue-twister effect cannot be due to visual confusions comes from the demonstration of this effect in Chinese, a nonalphabetic orthography (Zhang & Perfetti, 1993). Like their English-speaker counterparts, Chinese readers take longer to silently read tongue-twister stories compared to controls, and they make more errors in their recall. Finally, a functional MRI study of the silent reading of tongue twisters demonstrates that such sentences engage various phonological sites (e.g., Broca's area, the left angular/supramarginal gyri, and the areas along the left superior temporal sulcus) compared to control sentences (Keller et al., 2003). As these authors conclude, tongue twisters twist not only the tongue but also the brain – specifically, the regions involved in phonological processing and maintenance.

Why do people bother to decode the phonological forms of sentences, even though doing so interferes with the experimental task of digit recall? The answer to this puzzle becomes immediately apparent once we consider the properties of working memory. To escape oblivion, words must be maintained in a memory buffer called working memory (Baddeley, 1986). Working memory maintenance, however, is executed using a phonological format – try to memorize a phone number, and this will immediately become evident. So the phonological decoding of printed materials is mandated by the format of our short-term memory system. And because digit recall puts additional demands on working memory, it interferes with the phonological maintenance of sentences, an interference that is further exacerbated by the phonological similarity between words and digits.

12.3.2 Phonological activation in single-word recognition

The phonological format of working memory explains why all linguistic materials – printed or spoken – must ultimately undergo phonological encoding. Phonological decoding, however, begins immediately upon the recognition of single isolated words, even in tasks that impose only the slightest demands on working-memory maintenance. We now turn to examine the mechanisms mediating the phonological decoding of isolated words.

Alphabetic orthographies such as English allow readers to obtain phonological representations in two ways (see Figure 12.3). One method obtains the word's phonological form even before it is retrieved from the lexicon (i.e., pre-lexically) by mapping its graphemes to phonemes – a process known as *phonology assembly* (marked by the non-continuous line in Figure 12.3). English speakers, for example, know that the letter *d* signals the phoneme /d/, o can signal /ɔ/, and g corresponds to /g/. By relying on such regularities, readers can assemble the

C:/ITOOLS/WMS/CUP-NEW/3328241/WORKINGFOLDER/BERE/9780521769402C12.3D 291 [280–306] 28.8.2012 6:16AM

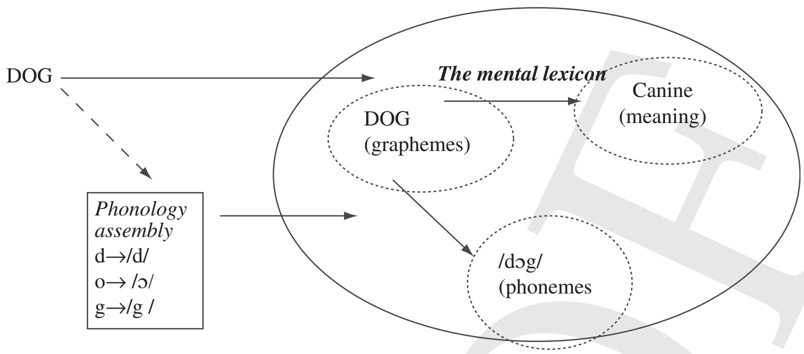

Figure 12.3 Two routes to phonology form print: assembled and addressed phonology

phonological representations of many printed words – familiar and novel. Familiar words such as *dog*, however, are also stored in the mental lexicon, linked to their phonological forms. Such words can thus be decoded along a second lexical route. The lexical route retrieves the word's lexical phonological representation directly from its graphemic form, a process known as *addressed phonology* (marked by the continuous line in Figure 12.3). While alphabetic systems such as English allow readers to obtain phonology from print along either the assembled or addressed route, in other writing systems – both alphabetic (e.g., Hebrew) and nonalphabetic (e.g., Chinese) – a word's phonological form cannot be fully generated by phonology assembly, so readers must rely on lexical retrieval to a greater extent. While the precise origins of a word's phonological form – assembled or addressed phonology – vary depending on the orthography, reading skill, and subtle properties of task demands, the reliance on some form of phonological encoding appears to occur universally, in all writing systems. This conclusion is supported by literally hundreds of published studies in many languages. Here, we will illustrate some of main findings from behavioral methods. Additional insights from neuroimaging studies are discussed in the next section.

12.3.2.1 Phonological predictability effects

A common method to gauge the contribution of phonology assembly in reading concerns predictability effects. This method exploits the well-known fact that, too often, phonological forms are only partly predictable from print. Consider English, for instance. While words like *dog* and *cat* can be reliably decoded by mapping their graphemes to phonemes, other words, like *come* and *put* are not fully predictable, as the assembly of their phonology would yield incorrect

C:/ITOOLS/WMS/CUP-NEW/3328241/WORKINGFOLDER/BERE/9780521769402C12.3D 292 [280–306] 28.8.2012 6:16AM

over-regularized forms, rhyming with *home* and *mute*. Such unpredictability can be annoying for readers, but it comes in handy to researchers who wish to determine how reading works. The rationale is simple: If English readers rely on phonology assembly, then phonologically unpredictable words (e.g., *come*) should exert a cost compared to predictable controls (e.g., *home*) – they should take longer to read and produce more errors. Such costs have indeed been detected by many studies (e.g., Andrews, 1982; Baron & Strawson, 1976; Glushko, 1979; Jared et al., 1990). While early research obtained these effects only for unfamiliar words, or unskilled readers (e.g., Seidenberg, 1985; Seidenberg et al., 1984), subsequent studies (Jared, 2002) documented predictability effects even for highly familiar words (provided that their pronunciation is highly unpredictable, i.e., that their incorrect pronunciation rhymes with many more familiar words than the correct one). The generality of these effects is important because it demonstrates that skilled readers assemble phonological forms to all words – rare or familiar.

Nonetheless, these findings are limited in several ways. First, predictability effects are typically observed only when people are asked to read the word aloud, but not in the silent lexical decision tasks (e.g., Berent, 1997; Seidenberg et al., 1984), a finding that is sometimes interpreted to suggest that silent reading is not mediated by phonology assembly. A second limitation of predictability effects is that they gauge reliance on phonology only when a word's pronunciation is unpredictable (e.g., for *come*). Accordingly, this method cannot determine the role of phonology when a word's phonological form is predictable – for phonologically "regular" words (e.g., for *home*) or in "transparent" alphabetic orthographies (e.g., Spanish, Italian) – and it is altogether useless in nonalphabetic orthographies (e.g., Chinese). To address these limitations, we now turn to a second marker of phonological decoding – homophony effects.

12.3.2.2 Homophony effects

Homophony in phonological priming/masking

Homophony effects gauge readers' sensitivity to phonological similarity – either the similarity among words (e.g., *rose*, *rows*) or nonwords (e.g., *rose–roze*). If phonological representations are employed in reading, then letter strings that share their phonology (e.g., *rose–rows)* should be perceived as similar, and this similarity should facilitate lexical access. To see why, let's think of lexical access as the opening of a door to the mental lexicon. If the lexical door has a phonological entry code, then words sharing the same phonological form (e.g., *rose*, *rows*, *roze*) should all be equally able to open the door. So once two homophones are presented in succession (e.g., *roze–rose*), *roze* will crack the lexical door of *rose*, and as *rose* follows, its identification (i.e., access to its lexical entry) should be easier.

C:/ITOOLS/WMS/CUP-NEW/3328241/WORKINGFOLDER/BERE/9780521769402C12.3D 293 [280–306] 28.8.2012 6:16AM

It is of course possible, however, that *roze* is helpful because it shares some of *rose*'s letters, rather than its sound. To control for this possibility, we can compare the facilitation from *roze* to *roge*, for instance. These two nonwords – *roze* and *roge* – are matched for their spelling similarity to *rose*, but differ on their phonological overlap. If the lexical entry to *rose* is specifically mediated phonology, then *roze* should have an advantage over *roge* – a case of phonological priming. If the process of extracting phonological representation from print occurs rapidly, then phonological priming should emerge even when the first word (a prime) is presented subliminally (see 10). In a similar fashion, presenting *roze* after *rose* should reinstate its phonological form, and consequently, it should facilitate its identification compared to spelling controls. As with priming, these effects, known as phonological masking, should obtain even when both words are displayed extremely briefly, masked by visual shapes.

(10) Phonological priming effects

Target:	rose
Phonological prime:	roze
Spelling control:	roge

These predictions have been amply supported in numerous orthographies. Research by Charles Perfetti and colleagues has demonstrated that the identification of English words benefits from a brief phonological prime or mask, presented for as little as 35 ms under heavy visual masking (Perfetti & Bell, 1991; Perfetti et al., 1988). The finding that phonological similarity can affect word identification even after a brief encounter suggests that these representations become available rather quickly. Moreover, these effects obtain for all words, both infrequent and highly familiar, suggesting once again that the contribution of phonological representations to reading is general in scope.

Phonological priming and masking effects have since been detected in numerous orthographies and writing systems. Among the alphabetic writing systems, phonological masking and priming have been reported not only in orthographies that readily support the assembly of phonology from print, such as French (e.g., Ferrand & Grainger, 1992) and Spanish (e.g., Carreiras et al., 2009), but also the opaque consonantal script of Hebrew (e.g., Berent & Frost, 1997; Frost et al., 2003). Moreover, phonological priming has been demonstrated even in nonalphabetic writing systems – in Japanese (e.g., Buchanan & Besner, 1993; Chen et al., 2007) and Chinese (e.g., Perfetti & Zhang, 1991; Tan & Perfetti, 1997). Unlike English, however, Chinese does not allow one to assemble phonology by mapping graphemes to phonemes on the fly, so phonological priming in Chinese reflects the retrieval of stored phonological forms from the lexicon (addressed phonology), rather than its online assembly.

C:/ITOOLS/WMS/CUP-NEW/3328241/WORKINGFOLDER/BERE/9780521769402C12.3D 294 [280–306] 28.8.2012 6:16AM

Although the phonological representations available to Chinese and English readers are obtained by different mechanisms, both orthographies are decoded by reliance on phonology.

Phonological interference effects

The findings reviewed so far suggest that the reflex of phonological decoding is quite robust. A true reflex, however, is judged by its automaticity. Once its triggering conditions are present, a reflex will proceed to completion, irrespective of whether its immediate effect is desired. The ultimate test of the phonological decoding reflex thus concerns not circumstances in which it is expected to help reading but those in which it is potentially deleterious. Our question is whether phonological decoding takes place under such conditions.

Semantic categorization experiments present one such case. Participants are asked to judge whether a word forms part of a semantic category (e.g., is *rose* a flower) – a task that clearly does not require that phonology be decoded from print. In fact, it tacitly discourages participants from doing so. Because many trials include foils that are phonologically similar to a flower exemplar (e.g., *rows*; *roze*), reliance on phonology is detrimental. But results show that readers nonetheless rely on phonology. Consequently, people tend to incorrectly accept homophone and pseudohomophone foils (e.g., categorize *roze* as a flower) relative to spelling controls, and these effects obtain in both alphabetic (English: Jared & Seidenberg, 1991; Van Orden, 1987; Van Orden et al., 1988; French: Ziegler et al., 1999) and nonalphabetic orthographies (Chinese: Perfetti & Zhang, 1995; Tan & Perfetti, 1997; Japanese Kanji: Wydell et al., 1993). People in these experiments do not fall for phonological foils on every trial (e.g., they do not invariably categorize *roze* as a flower) because they can put their phonological instincts in check using a spell-verification mechanism. But if phonology is nonetheless active, homophones should be more likely to slip in, which is precisely what is observed.

In fact, people demonstrably rely on phonological representations even when they are explicitly required to avoid reading altogether, in the Stroop task. The Stroop task, discussed in previous chapters, presents participants with letter strings printed in colors (e.g., the word *green* printed in the color red), and participants are asked to name the color of the ink (*red*) while ignoring the printed letter. The classical finding is that readers automatically decode the words, and consequently, color naming is impaired when the printed word spells the name of an incongruent color (Stroop, 1935). Subsequent research, however, showed that Stroop interference obtains not only from the conventional spelling of color names but also by their homophones (e.g., the word *bloo* printed in red). These homophonic Stroop effects have been reported in various alphabetic systems (English: Naish, 1980; French: Ziegler et al., 1999; Hebrew: Berent et al., 2006; Tzelgov et al., 1996) as well as in Chinese (Guo et al., 2005; Spinks et al., 2000; Wang et al., 2010). Such demonstrations suggest that the encounter with a printed

C:/ITOOLS/WMS/CUP-NEW/3328241/WORKINGFOLDER/BERE/9780521769402C12.3D 295 [280–306] 28.8.2012 6:16AM

word automatically triggers phonological decoding even when people attempt to avoid doing so.

12.4 Reading recruits the phonological brain network

Another test of the role of phonological reflexes in reading is presented by the brain networks that mediate silent reading. If silent reading recruits the phonological system, then printed words should engage brain mechanisms implicated in the phonological processing of spoken language. The previous section has already offered one illustration of the role of phonological brain networks in the silent reading of tongue-twister sentences. In what follows, we further evaluate this prediction at the level of single words.

In one study (Buchsbaum et al., 2005), participants were presented with pronounceable monosyllabic pseudowords across two modalities – either in print or aurally. Results revealed numerous brain regions that were common to the two conditions, including the left planum temporale, left STG, and the left middle temporal gyrus, as well as the left IFG – regions that correspond to the phonological brain network discussed in Chapter 10.

Further evidence for the phonological functions of these temporal-lobe sites is offered by their differential contribution to the naming of novel letter strings – either meaningless letters strings (i.e., pseudowords, e.g., *blag*) or ones homophonous to real words (e.g., *burth*) – as compared with irregular words (e.g., *pint*). Because novel words are not lexically stored, their phonological form can be obtained productively only by mapping graphemes to phonemes (i.e., phonology assembly). Irregular words, by contrast, have a phonological form that is unpredictable from print, so their pronunciation requires lexical retrieval (i.e., addressed phonology). To the extent that assembled and addressed phonology engage different brain regions, one would thus expect novel words and irregular words to elicit different patterns of activation. An fMRI experiment by Simos and colleagues supports these predictions (Simos et al., 2002). Irregular words activated the posterior middle temporal gyrus and the middle temporal lobe to a greater extent than novel words, suggesting that those regions specifically mediate the retrieval of addressed phonological forms. In contrast, activation in the posterior STG correlated with the pronunciation time of nonwords (but not exception words), suggesting that this region subserves the assembly of phonology from print. Interestingly, however, the posterior STG also contributed to the processing of irregular words, indicating that naming invariably triggers phonology assembly of all words – regular or irregular. A subsequent magnetoencephalography study (Simos et al., 2009) observed bilateral activation of the posterior superior temporal gyri and inferior frontal gyri using the lexical decision task. These findings are significant because they suggest that the phonological network mediates silent reading, and it is active quite generally,

C:/ITOOLS/WMS/CUP-NEW/3328241/WORKINGFOLDER/BERE/9780521769402C12.3D 296 [280–306] 28.8.2012 6:16AM

irrespective of the frequency of those words. Moreover, a meta-analysis of imaging studies of reading across Western alphabetic and Eastern orthographies (Bolger et al., 2005) concluded that the left posterior STG (Brodmann Area 22) is active in all orthographies, including the nonalphabetic Chinese and Japanese Kanji systems (but see Liu et al., 2009; Siok et al., 2008, for different conclusions). These studies illustrate the conclusions emerging from a very large literature that links the phonological network of spoken language to reading. These findings underscore an interesting link between mind and brain. Just as the cognitive architecture of reading is based on older phonological principles, so does the phonological brain assemble the reading networks by recycling older substrates that mediate core phonology (Dehaene et al., 2010; Dehaene & Cohen, 2011).

12.5 Grammatical phonological reflexes in reading

The behavioral and neural evidence reviewed in previous sections make it clear that reading entails phonological decoding. It is the recovery of phonological structure that allows readers to determine that two printed words are homophonous (e.g., *rose–rows–roze*) and leads to the over-regularization errors of irregular words (e.g., *come*). While there results firmly establish that readers extract some phonological representations, they do not determine whether these representations are in fact identical to the ones extracted in spoken language processing. To determine that two words are homophonous, for example, readers could rely on simple phonological codes that list only phonemes and their linear order. This possibility would entail a rather superficial link between reading and the putative system of core phonology – while both systems might encode phonological information, the representations computed by the two systems would be qualitatively different. On an alternative account, the representations computed to visual and spoken language are isomorphic: They encode the same set of primitives, and they are constrained by the same set of grammatical constraints. The link between reading and core phonology is thus intimate. The neural evidence summarized in the previous section is consistent with this latter possibility. Further support for this view is presented by behavioral evidence showing that silent reading yields structured phonological representations. These representations specify the same phonological primitives implicated in spoken language processing, and they are subject to some of the same grammatical constraints – both language-particular and universal.

12.5.1 Phonological primitives in silent reading

The review of spoken language phonology suggests that phonological features, CV frames, and syllables form the representational primitives of core phonology. Each of these elements has also been implicated in reading.

C:/ITOOLS/WMS/CUP-NEW/3328241/WORKINGFOLDER/BERE/9780521769402C12.3D 297 [280–306] 28.8.2012 6:16AM

Consider, for example, the role of phonological features in silent reading. If the representations assembled in reading encode the feature composition of segments, then words that share most of their features should be perceived as phonologically similar even if they do not share phonemes. In line with this prediction, printed words that share most of their phonological features (e.g., *sea–zea*, which differ only on the voicing) have been shown to prime each other to a greater extent than controls (e.g., *sea–vea*, which differ by voicing and place of articulation; Lukatela et al., 2001; see 11). Remarkably, readers are sensitive to phonological features even when the task, lexical decision (e.g., is *sea* a real English word?), does not require phonological encoding or articulatory response. Additional auxiliary analyses rule out the possibility that the similarity between *sea* and *zea* reflects visual features. Together, these findings suggest that the computation of sub-segmental phonological detail might be quite general.

(11) Priming by phonological features in silent reading

Target:	sea
Phonological similar prime:	zea
Control prime:	vea

Other findings indicate that such sub-segmental information includes even non-contrastive features. It is well known that the acoustic duration of the vowel in CVC words differs as a function of the following coda consonant. Vowels followed by voiced codas (e.g., *pl**ea**d*) are longer than ones preceding voiceless codas (e.g., *pl**ea**t*). This acoustic difference is analog and non-contrastive (i.e., no English words differ solely on this dimension), but it nonetheless reflects a regular characteristic of phonological systems. Strikingly, however, this subtle phonetic contrast has been shown to affect silent reading: Readers take longer to classify *plead* as an English word compared to its counterparts *pleat*, even though the materials are printed and the task (lexical decision) requires no articulatory response. This effect, first demonstrated by Abramson and Goldinger (1997), was subsequently replicated in both behavioral (Ashby et al., 2009; Lukatela et al., 2004) and EEG measures (Ashby et al., 2009).

Moving up in the hierarchy of phonological primitives, readers encode various prosodic constituents. Readers in different orthographies are sensitive to the syllable structure of multisyllabic words. Participants in these experiments are presented with a printed multisyllabic target word beginning with either a CV or a CVC syllable (e.g., *ba.sin* vs. *bas.ket*; see 12), and each such target is preceded by either a CV prime (e.g., *ba*) or a CVC prime (e.g., *bas*). Results show that readers are sensitive to the congruency between the syllable structure of the target and prime. For example, people are faster to classify *basket* as an English word when preceded by *bas* compared to *ba* despite not being required to articulate either the

PROOF

target or prime. Syllable congruency effects, moreover, obtain not only in so-called syllable-timed languages – languages like Spanish and French (Carreiras et al., 2005; Chetail & Mathey, 2009; Colé et al., 1999), where syllable structure is well defined – but even in English, a stressed-time language whose syllable structure is relatively opaque. In a series of examples, Jane Ashby and colleagues demonstrated that the brief priming of the target by a syllable-congruent prime that is presented parafoveally (i.e., in the area surrounding the fovea, such that readers are unaware of the prime) facilitates target recognition (Ashby & Rayner, 2004) and decreases the negative N1 evoked-potential brain response to the target word (Ashby, 2010; Ashby & Martin, 2008).

(12) Syllable priming effects in silent reading

		Target	
		CV (e.g., *ba.sin*)	CVC (e.g., *bas.ket*)
Prime	Congruent	*ba*	*bas*
	Incongruent	*bas*	*ba*

12.5.2 Grammatical constraints in silent reading

Not only do the readers decode phonological primitives from print, but they further subject these representations to the same grammatical constraints operative in spoken language. The evidence is particularly strong when it comes from novel words. Unlike familiar words, the phonological forms of novel words cannot be retrieved from the lexicon, so such phonological effects imply a productive grammatical mechanism that is operative online in reading. Previous chapters have discussed various examples of such grammatical constraints in detail, so here we will review some of these examples rather briefly.

Recall, for example, that Hebrew constrains the location of identical segments in the root – roots with initial identical consonants are ill formed (e.g., *ssm*) whereas roots with identical consonants at their end are quite frequent (e.g., *smm*). A second restriction bans non-identical segments that share the same place of articulation (i.e., homorganic consonants, such as the two labials in *smb*). Results show that Hebrew speakers are sensitive to both constraints in silent reading (Berent et al., 2001b; Berent et al., 2004). Because ill-formed strings are less wordlike, such strings should be easier to identify as nonwords. And indeed, words derived from ill-formed roots (e.g., *ssm*, *smb*) are identified more readily than well-formed controls (e.g., *smm*), matched for segment co-occurrence. In fact, the restriction on root structure modulates reading even when people are explicitly asked to avoid reading altogether, in a modified Stroop task. In these experiments, Hebrew speakers were presented with words printed in color. As in the typical Stroop task, the task was to name the color of

C:/ITOOLS/WMS/CUP-NEW/3328241/WORKINGFOLDER/BERE/9780521769402C12.3D 299 [280–306] 28.8.2012 6:16AM

the ink while ignoring the printed letters. These letters, however, corresponded to novel words constructed with novel roots – either well-formed roots (e.g., *smm*) or ill-formed controls (e.g., *ssm*). Because ill-formed roots engage the grammar to a lesser extent than better-formed controls, readers should find it easier to ignore ill-formed structure, a fact that might leave them free to quickly name the color and ignore the root. Results indeed showed that the structure of the root affected color naming: People named the color faster with ill-formed *ssm*-type roots compared to well-formed controls (Berent et al., 2006). These results indicate that grammatical phonological knowledge concerning Hebrew phonotactics constrains reading automatically, in a reflex-like manner.

Additional results from English suggest that these grammatical constraints might include principles that are potentially universal. One such case concerns the markedness of syllable frames. Across languages, syllables that manifest an onset are preferred to onsetless ones, and simple syllable margins are preferred to complex ones. For these reasons, syllables manifesting an onset and simple margins (e.g., CVC) are unmarked compared to syllables with no onset and complex margins (e.g., VCC). If marked structures are less likely to engage the language system, then VCC structures should be easier to ignore than CVC ones. The results from modified Stroop experiments are consistent with this prediction (Marom & Berent, 2010). In these studies, people were presented with novel words printed in color. None of these words shared phonemes or graphemes with the color name, but in some cases, the color name and the word shared the same CV skeletal frame (e.g., the word TROP printed in the color *black* – a CCVC frame), whereas in others they were incongruent. Crucially, incongruent frames were either relatively unmarked (e.g., the CVC frame in GUF) or more marked (e.g., VCC, as in OCP). As expected, people were sensitive to the skeletal congruency between the color and the word, resulting in faster naming time in the CV-congruent condition (see also Berent & Marom, 2005). But crucially, responses to the incongruent condition were modulated by markedness, such that incongruent frames were easier to ignore when they were marked (e.g., for VCC compared to CVC).

Other, putatively universal markedness restrictions concern sonority sequencing. Recall that speakers constrain the sonority distance of auditory onsets that are unattested in their language – the smaller the distance, the more likely is the onset to be epenthetically repaired. In particular, people tend to misidentify ill-formed onsets of falling sonority as identical to their epenthetic counterparts (e.g., *lbif*=*lebif*), they are less likely to do so for sonority plateaus (e.g., *bdif*), and they are least likely to misidentify well-formed onsets of rising sonority (e.g., *bnif*). Interestingly, these sonority restrictions extend to printed words: CCVC syllables with ill-formed onsets are harder to distinguish from their CəCVC counterparts (Berent & Lennertz, 2010; Berent et al., 2009). Although the findings from printed materials are less robust than those with auditory stimuli,

C:/ITOOLS/WMS/CUP-NEW/3328241/WORKINGFOLDER/BERE/9780521769402C12.3D 300 [280–306] 28.8.2012 6:16AM

and they obtain only when the task allows for sufficient detailed processing of the printed words (Berent, 2008), the available results nonetheless show that, given sufficient processing time, silent reading is shaped by grammatical constraints and yields phonological forms that are largely isomorphic to the ones extracted from spoken language.

12.5.3 Dyslexia as a phonological disorder

The previous sections have demonstrated that reading ability is intimately linked to phonological competence. Not only does phonology form the basis for the design of writing system, but it is routinely recruited in its online decoding, in reading single words and texts. The close link between reading ability and core phonology carries some direct implication to reading disability as well. If skilled reading relies on core phonology, then deficits to core phonology are expected to impair the acquisition of reading skill. In what follows, I evaluate this prediction by examining the phonological competence of individuals with developmental dyslexia and evaluate the etiology of this disorder.

12.5.3.1 Phonological deficits in dyslexia

Developmental dyslexia is a deficit in the acquisition of reading skill that is unexplained by intelligence, emotional, motivation, and social factors, and it affects between 5–17 percent of the population (Shaywitz, 1998). Dyslexia is a complex disorder with multiple causes, including visual deficits, working memory and attention limitations. Many dyslexic individuals, however, also demonstrate subtle phonological impairments (Dehaene, 2009; Ramus, 2001; Shaywitz, 1998).

One aspect of this impairment is evident in reading itself, especially when it comes to the decoding of novel words (e.g., *blig*). Novel words indeed exert far greater phonological demands than existing words. Existing words (e.g., *block*) have memorized phonological forms that can be retrieved by association with their spellings, akin to the retrieval of a person's name from the sight of his or her face. By contrast, novel words' pronunciations can only be obtained "from scratch" – by a productive process that maps each of their grapheme to phonemes (e.g., b➔/b/). And before a child can even learn the mapping, he or she must first become aware that words comprise phonemes (Liberman, 1989). A failure to encode the phonological structure of spoken language, to gain awareness of its constituent phonemes, and to automatically link them to graphemes is bound to impair nonword naming. And indeed, many studies have shown that unskilled readers of alphabetic orthographies manifest difficulties in the decoding of non-words (e.g., Paulesu et al., 2001; Rack et al., 1992; Ziegler et al., 2009).

But the phonological difficulties of dyslexics are not limited to reading tasks nor are they confined to alphabetic writing systems. Numerous studies have

C:/ITOOLS/WMS/CUP-NEW/3328241/WORKINGFOLDER/BERE/9780521769402C12.3D 301 [280–306] 28.8.2012 6:16AM

shown that dyslexics experience greater difficulties in the processing of spoken language. For example, dyslexics are impaired in tasks that elicit explicit analysis and awareness of the phonological structure of spoken language – tasks such as rhyming (e.g., *does* bat *rhyme with* hat?) phoneme deletion (say *block* without the first sound), and spoonerisms (e.g., *bed+lot➔leg+bot)*. Such deficits have been repeatedly demonstrated not only in alphabetic writing systems (Bishop & Snowling, 2004; Manis et al., 1997) but also in Chinese (Siok et al., 2008). Although the failure to develop phonemic awareness could also result from illiteracy (Morais et al., 1979), and as such, its impairment in dyslexia could reflect a symptom of the disorder, rather than its cause – reading skill alone is insufficient to explain the phonological delays of dyslexics. Indeed, the phonological deficits of dyslexics are evident even when compared to controls matched for reading skill (Ziegler et al., 2009), and they even extend to tasks that do not require awareness of phonological structure – in repeating orally presented words, rapidly naming objects, and the maintenance of words and digits in memory (Paulesu et al., 2001; Ziegler et al., 2009). These deficits, moreover, can be traced to difficulties in extracting phonological features from speech signal, such as place (e.g., *ba* vs. *da*) and manner of articulation (e.g., *ba* vs. *fa*; e.g., Mody et al., 1997; Serniclaes et al., 2001; Ziegler et al., 2009). Because many dyslexics do not manifest frank linguistic deficits in either comprehension or production, such subtle speech perception deficits are typically detectable only when the perception of speech is made more challenging, by degrading the speech signal, presenting it masked in noise (Ziegler et al., 2009), or eliciting discrimination among sounds presented in rapid succession (Mody et al., 1997). Nonetheless, these abnormalities in the perception of speech sounds have been documented in several studies. These findings indicate that the phonetic representations extracted by dyslexic individuals from speech are fragile. In fact, these deficits in speech perception can be detected already in infancy.

Longitudinal studies have shown that event-related brain potential responses to speech stimuli obtained from an individual at infancy can be linked to that individual's reading scores at 8 years of age (Molfese, 2000). Moreover, infants from families with high incidence of dyslexia differ from controls in their perception of phonetic contrasts, such as the Finnish contrast in consonant duration (e.g., the contrast between *ata* and *atta*; Leppänen et al., 2002) and the Dutch contrast between /bak/ and /dak/ (van Leeuwen et al., 2007). Specifically, while left hemisphere brain responses of control infants differentiated /bAk/ and /dAk/, the brain responses of 2-month-old infants at high-risk of dyslexia failed to differentiate among these exemplars, and their brain activity originated predominantly in the right hemisphere.

The fact that familial pedigree presents a risk factor for dyslexia also underscores the strong hereditary basis of this disorder. It has long been noted that dyslexia runs in families. Families, of course, share both genes and environment,

C:/ITOOLS/WMS/CUP-NEW/3328241/WORKINGFOLDER/BERE/9780521769402C12.3D 302 [280–306] 28.8.2012 6:16AM

so familial patterns do not uniquely demonstrate a genetic link. Twin studies, however, allow one to dissociate genetic from environmental factors by comparing the prevalence of the disorder among monozygotic twins – twins who share all their genes – and dizygotic twins who share on average 50 percent of their genes. To the extent that dyslexia has a genetic basis, one would expect it to be more prevalent among monozygotic twins. Twin studies indeed show that if one sibling suffers from dyslexia, his or her twin is significantly more likely to exhibit this disorder when the twins are monozygotic. The precise heritability of dyslexia (i.e., the amount of variance associated with genetic factors) varies across studies (Bishop & Snowling, 2004), but according to one estimate (Pennington & Bishop, 2009) the overall heritability of dyslexia is 0.58 (Pennington & Bishop, 2009), and the heritability that is specifically associated with phonological awareness reaches 0.53 (Byrne et al., 2002). Subsequent genetic studies have identified several candidate genes, including *DYX1C1*, *DCDC2*, *KIAA0319*, *ROBO1* (Galaburda et al., 1985; Pennington & Bishop, 2009; Shastry, 2007).

Although there is strong evidence that individuals with dyslexia suffer from a highly heritable deficit to speech processing, these data do not establish its source – whether it originates from lower-level impairment to phonetic processing or whether it extends to the phonological grammar. Although there is a large literature on the perception of phonetic categories, we know very little about the sensitivity of dyslexics to phonological structure, and the existing findings are inconsistent. Some researchers found that dyslexics exhibit phonotactic deficits. For example, dyslexics are less sensitive to phonotactic probability (Bonte et al., 2007), and they experience difficulties in the production of consonant clusters in unstressed syllables (Marshall & Van Der Lely, 2009). Other researchers, however, failed to find any phonotactic deficit (Szenkovits et al., 2011). Like normal (French) controls, dyslexic individuals were sensitive to the contrast between onsets such as *bl* (attested in French) and *dl* (which is unattested), and both groups showed a similar tendency to misidentify unattested onsets as their licit counterparts (e.g.,*dl*➔ *bl*). Similarly, dyslexics distinguished between phonological processes that are obligatory in their language and ones that are unattested. They correctly produced voicing assimilation, a process that is obligatory in their language (e.g., *cape gris* ➔ [kabgʁiz]), but did not assimilate place of articulation – a process that is unattested in French (*zone portuaire*➔* [zompoʁtyɛʁ]), and like their typical counterparts, they tended to perceptually compensate for assimilation (i.e., they failed to detect assimilation in legal contexts) only when the assimilatory process was legal in their language – for voicing, but not place assimilation.

The small number of studies makes it difficult to draw any firm conclusions regarding the status of the phonological grammar in dyslexia. One possibility is that at least some dyslexics manifest deficit in the phonological grammar, even if they do not otherwise show any frank speech or language disorder. The failure

C:/ITOOLS/WMS/CUP-NEW/3328241/WORKINGFOLDER/BERE/9780521769402C12.3D 303 [280–306] 28.8.2012 6:16AM

to detect such impairments consistently might be due to individual differences in the severity of the deficit, task sensitivity, or the specific aspect of phonological competence under investigation. Indeed, existing studies have tested the sensitivity of dyslexics only to structures attested in their language, so in such cases, individuals could compensate for their grammatical deficit by relying on lexical memory. Unattested structures, however, might be less likely to benefit from lexical compensation, so it is conceivable that future investigations of such cases might reveal a grammatical deficit. On an alternative account, the phonological deficit of dyslexics originates from low-level difficulties in the extraction of phonetic structure, rather than the phonological grammar per se. Because imprecise phonetic representations can lead to imprecise phonological forms, such low-level deficits could occasionally interfere with tasks that require phonological judgment (e.g., phonotactic sensitivity) even if the phonological grammar is otherwise intact. It is indeed well-known that dyslexic individuals manifest deficits in the extraction of phonetic features, and existing studies have not ruled out the possibility that phonological errors might result from phonetic processing impairments (Bonte et al., 2007; Marshall & Van Der Lely, 2009). The precise locus of the phonological deficits in dyslexia awaits further research.

12.5.3.2 The etiology of dyslexia

While existing research makes it clear that many individuals with dyslexia manifest phonological deficits, the nature of these impairments remains unknown. These outstanding questions regarding the phonological deficit in dyslexia also make it difficult to evaluate the etiology of the disorder. As with other language developmental disorders, such as Specific Language Impairment and Speech Sound Disorder – hereditary disorders that exhibit high comorbidity with dyslexia (Pennington & Bishop, 2009) – the class of potential explanations for dyslexia ranges from domain-specific accounts to domain-generalist explanations. Domain-specific accounts attribute these disorders to a specialized language system, either a specialized speech perception mechanism (e.g., Mody et al., 1997) or components of the grammar (van der Lely, 2005; van der Lely et al., 2004). Alternative explanations view those linguistic deficits as secondary to basic impairments in either nonlinguistic systems (e.g., the magoncellular system; Stein & Walsh, 1997) or domain-general mechanisms, such as procedural learning (Ullman & Pierpont, 2005), lexical retrieval (Wolf et al., 1986), and working memory (Gathercole & Baddeley, 1990). Of these various domain-general accounts of dyslexia, low-level auditory processing deficits have received wide attention. In a series of influential studies (Benasich & Tallal, 2002; Tallal, 2004; Tallal & Piercy, 1973; Temple et al., 2000; Temple et al., 2003), Paula Tallal and colleagues proposed that dyslexia and specific language impairments result from low-level deficits in the processing of brief or rapid auditory events. In support of

PROOF

C:/ITOOLS/WMS/CUP-NEW/3328241/WORKINGFOLDER/BERE/9780521769402C12.3D 304 [280–306] 28.8.2012 6:16AM

this proposal, Tallal and colleagues demonstrated that the processing deficits of individuals with language-learning impairments extend to nonlinguistic stimuli – to auditory tones presented either briefly or in rapid succession – that these deficits are present in infants at high risk of language disorders (Benasich & Tallal, 2002), and that training on the discrimination of such brief auditory and linguistic events (specifically, on modified speech designed to slow or amplify rapid frequency transitions) might improve reading skill (Temple et al., 2003). Although deficits to the processing of brief/rapid auditory events are not seen in all adult dyslexics (e.g., Mody et al., 1997), they appear to be more prevalent in younger children (Tallal, 2004).

The auditory origins of dyslexia and their transitory developmental nature are both captured by an influential genetic model of this disorder. Examining the brains of deceased individuals with dyslexia, Albert Galaburda has noticed subtle cortical anomalies that originate from a disruption to neural migration during embryonic development (Galaburda et al., 1985). Indeed, several of the candidate susceptibility genes for dyslexia have been linked to neural migration and growth (Galaburda et al., 2006), and an animal rat model (e.g., Burbridge et al., 2008) shows that disruption to these genes results in cortical and subcortical anomalies that mirror the ones found in dyslexic individuals, and it also yields similar behavioral symptoms. Like dyslexic individuals, affected rats manifest disruption to the discrimination of auditory tones presented in rapid succession (Peiffer et al., 2004), these deficits are larger in juvenile animals (Peiffer et al., 2004), and they are more prevalent in males (Galaburda et al., 2006), a finding that mirrors the larger prevalence of dyslexia in human males (Rutter et al., 2004). These results open up the possibility that the phonological deficits of dyslexics might originate from prenatal insults to brain development that result in impairments in the perception of rapid/brief auditory stimuli. While these auditory deficits might be initially present in all individuals, they might eventually ameliorate in later development (Galaburda et al., 2006).

The neural migration theory is unique in its ability to offer a comprehensive account that ranges the entire gamut from genes to behavior: It links phonological deficits to specific brain abnormalities, identifies genetic mutations that cause those neural changes, and traces the dynamic unfolding of these gene–brain–behavior interactions throughout development. It is still too early to tell whether this model can in fact capture the full range of behavioral and neural abnormalities reported in humans, but the available data are promising. At first blush, this proposal would seem to challenge the core phonology hypothesis. If hereditary phonological deficits could originate from low-level auditory disorders, then perhaps auditory mechanisms might be sufficient to account for phonological competence in healthy individuals. But the challenge is only apparent. To begin with, it is unclear whether most dyslexics exhibit a deficit to the phonological grammar, so their putative auditory processing impairment could be unrelated to phonological competence (Berent et al., 2012c; Ramus & Szenkovits, 2006). Even if the phonological grammar were impaired, this would hardly imply that auditory processing subsumes core phonology. Correlations between phonological and auditory processing could reflect some third common factor that governs the development or use of both systems. And even if these two systems share some of their hardware resources (brain sites and genes), they might still be segregated at the functional level – a conclusion defended in Chapter 10.

In summary, there is mounting evidence that dyslexia is associated with a host of phonological deficits, but many questions remain regarding the scope of these deficits and their origins. Concerning the scope of the disorder, we do not currently know whether most dyslexics manifest deficits in the organization of the phonological grammar, or whether their difficulties in the processing of spoken language are confined to the phonetic and acoustic levels. Similarly, it is unknown whether the phonological processing deficits of dyslexics are secondary to a lower-level auditory deficit, or whether they originate from injuries to the language system occurring at either the phonetic or the phonological level. So while the clear phonological deficit in dyslexia certainly underscores the link between normal reading and the phonological system, it is unclear whether the subtle deficits to spoken language processing that are characteristic of many dyslexics do in fact reside specifically in the grammatical phonological component.

12.6 Conclusion

Humans are equipped with several systems of core knowledge. Core knowledge systems, such as primitive number systems and naïve physics, each manifests unique design that is universal, adaptive, and present early in birth. Unlike the universal instinctive core knowledge systems present in infancy, other aspects of knowledge are the domain of the select few, they are discovered only in late development or adulthood through intense reasoning or scientific methods, and transmitted by explicit instruction. But the structure of those invented systems can be traced back to their origins in core knowledge systems. The propensity of those core knowledge systems to lay the foundation for cultural inventions is one of their defining features.

Previous chapters have examined the hypothesis that the phonological grammar forms a system of core knowledge. In support of this possibility, we have demonstrated that the system manifests universal substantive constraints that appear to be adaptive and active early in life. Here, we have seen that, like the core systems of number and physics, the putative core system of phonology lays the foundation to the invented systems of readings and writing.

Considering writing, we have seen that all full writing systems encode phonological primitives, and that the link between core phonology and writing

C:/ITOOLS/WMS/CUP-NEW/3328241/WORKINGFOLDER/BERE/9780521769402C12.3D 306 [280–306] 28.8.2012 6:16AM

recapitulates spontaneously in the writing systems invented by young children. While writing invariably encodes phonological structure, reading universally decodes it. The precise method and grain-size of decoding varies – transparent alphabetic writing systems like Italian and Spanish allow for a rapid extraction of phonology by mapping graphemes to phonemes; at the other extreme, the Chinese syllabic orthography conveys segmental phonology only at the syllabic levels. But despite those differences, in all writing systems, skilled readers extract phonological structure from print, and at least in alphabetic orthographies, they do so by relying on brain sites that mediate the phonological processing of spoken language. But the computation of phonological structure from print goes beyond the bare minimum of segmental phonology – a closer inspection offers numerous demonstrations that the representations computed on printed language exhibit significant overlap with those computed on spoken language, including shared primitives and structural constraints. And indeed, deficits in processing of spoken language are associated with dyslexia. While the precise source of the deficits – whether they result from a deficit to the language/speech system or to auditory processing – and their extent – whether they are confined to feature extraction or extend to the organization of the grammar – remains debated, the centrality of phonology to the design of writing and to its online decoding underscores the intimate link between these cultural technologies and the phonological system at their core.

PROOF