A componential approach for bilingual reading and comparative writing system research: The role of phonology in Chinese writing as a test case

Norbert Francis
Northern Arizona University

Abstract

The special circumstances of bilingual and second language literacy learning offer investigators an important additional vantage point from which to better understand the components of reading ability. Cross-writing system comparisons complement this perspective. Comparing writing systems, and how children learn to read through the medium of each system provides for tests of a number of hypotheses currently under discussion. One particularly instructive series of tests involves the contrast between alphabetic and non-alphabetic writing systems. This review of the research will examine proposals related to the role of phonology in word identification with a special focus on the morpho-syllabic/logographic Chinese orthography. A componential, or modular, approach to the study of reading ability will be evaluated in relation to claims made from different perspectives on the question of the activation of phonological representations in reading. In particular, is the Universal Phonological Principle, proposed by C. Perfetti, compatible with a modular approach to the study of reading ability?

Introduction

Research on cross-language comparisons provides special opportunities and vantage points for a better understanding of the components of reading ability. These comparisons can approach the analysis of skills and abilities from two different perspectives: the study of reading development in bilinguals (including here the subset of individuals who are second language learners), and the comparative study of reading across language and writing systems among monolingual literacy learners. An assumption that ties these two approaches together conceptually is that the research in reading in any single language is relevant in every way to second language and bilingual literacy. The same underlying cognitive structures form the foundation for reading ability. Conversely, the findings from bilingual literacy will help us better understand the components of reading ability in general. For beginning second language readers, imbalances and uneven development provide opportunities to examine the relevant components and sub-components.

The review of the research will begin with: (1) a brief overview of theoretical approaches to the study of reading that have made reference to the concept of modularity, followed by (2) a discussion of the current state of investigations in the field of bilingual and second language reading, giving examples of how a componential approach might be relevant to a number of ongoing research problems. Cross-language and cross-writing system comparisons will present us in the concluding sections (3) and (4) with a complementary viewpoint on the same issues, in particular comparisons that
involve non-alphabetic writing systems. Central to all of the above will be an assessment of the Universal Phonological Principle (UPP) and the related Universal Writing System Constraint (Perfetti, 2003; Perfetti & Dunlap, 2008) and how they might be understood in relation to one or another version of the modularity hypothesis. The important claim of the UPP is that the activation of phonology in (silent) reading cannot be by-passed in the use of any writing system.

The assessment of these models of reading is the primary objective of this review. The problem of determining which aspects of literacy are universal and which are specific to each writing system is still not well understood; and much work is still needed to even conceptualize all the critical research questions coherently. The last two sections will give readers a glimpse into this difficult task. As we will see, the study of the Chinese writing system, how it differs from alphabetic systems, and how it does not, holds one of the most important keys to this problem. For discussion purposes, an overview of the general approach of componential analysis, associated with the notion of modularity, will be presented as one way to frame the relevant issues. The specific purpose of this overview is to set the stage for a proposal: if future research shows that the Universal Phonological Principle and the Universal Writing System Constraint are correct, this result would be consistent with a modular approach to the study of reading. On the other hand, if Perfetti's theories cannot stand up to future disconfirming evidence, this would seriously call into question central pillars of the modularity hypothesis, and not just in the area of literacy. As such, the problems of bilingual literacy learning and cross-writing system comparison (in particular, the contrast between alphabetic and morpho-syllabic Chinese systems) present a critical test for an important debate in cognitive science.

The componential approach
There are a number of ways in which reading ability might be analyzed componentially; and the concept of modularity has been used to frame the various problems that arise in this type of analysis. This review of the research will do the same. While certain versions of the modularity thesis remain controversial, in particular those among them that we might characterize as “strong bottom-up” versions, it is important to keep in mind that there exists a broad diversity of views among researchers who favor one or another componential/modular approach to the study of language ability. Even if historically the concept traces its roots to Universal Grammar-oriented theories, there is no reason for insisting that the general idea of modularity is incompatible in every way with other schools of cognitive science. Keeping this in mind might turn out to be especially useful in the research on a complex ability such as reading. This section will provide an overview of a modular approach to the study of reading not for the purpose of making the claim that it is the most correct or comprehensive framework for literacy research. Rather, a more narrow claim will be offered for discussion: The main idea underlying the UPP, that activation of phonology in (silent) reading cannot be by-passed, even in non-alphabetic writing systems, follows from a componential/modular perspective on both the architecture of
language and literacy knowledge and the processing of written language. While Perfetti has not explicitly drawn this connection, the questions related to how linguistic knowledge structures interact in the deployment of an ability like reading are central to this claim. If his hypotheses prove to be incorrect, this would spell trouble for modularity (for strong and weak versions alike). Thus, current and future research evidence that is inconsistent with the UPP would lend support to holistic and integrativist theories of literacy (e.g., “whole language” theories).

Issues related to the concept of modularity come up in the discussion of a number of important questions in the research on literacy and bilingualism, for example:

(1) Can a useful distinction be made between explicit knowledge and metalinguistic awareness on the one hand and implicit knowledge on the other?

(2) In reading, how do the different subsystems of grammar come on-line and how do they interact with the other (non-linguistic) knowledge structures and processing mechanisms that are specific to reading? In relation to bilingual readers, this topic is related to the question of the differentiation between first language (L1) and second language (L2) systems:

(3) What are the essential points of contention in the debate opposing holistic social-constructivist theories and those that emphasize the importance of mastering the sub-skills of word identification? and

(4) How do we understand the concepts of pathway, circuit, and network, and might they be compatible with models that conceive of components? In this regard, what might be some points of contact between connectionist and modular approaches to the study of reading ability?

Let us take each of these points in turn, beginning with (1) above. A basic assumption underlying modular approaches is that not all knowledge of language, for example, is of the same kind. Especially when considering complex abilities, an important distinction is between metalinguistic explicit knowledge and implicit knowledge. The phonological level, being the most basic for reading, appears to reveal this distinction most clearly. Arguably, the phonological competence that young children attain in their L1 depends on no kind of deliberate attention to form, negative evidence, or application of deliberate learning strategy. From this point of view, the early spontaneous and universal emergence of phonological competence would be the hallmark of an innately pre-programmed cognitive specialization (Petitto, 2007). Assessments of metaphonological knowledge show evidence of developmental trends that are not spontaneous and universal, correlating (as implicit phonological competence does not) with literacy learning (Castles & Coltheart, 2004; Hu, 2004; Liberman, 1999; Morais & Kolinski, 2001).

The question of how the subsystems of language are deployed and how they interact with other cognitive domains in reading (question 2) assumes that reading ability can be analyzed (“fractionated”) into subsystems or autonomous components in the first place (Stanovich, 2003). To clarify terms here, “interaction” and “interactive models” imply something very different from “integration.” An “integrativist” model would reject all versions of modularity,
while interaction assumes connectivity and interface among components. In fact, any conception of a modular/componential architecture that omits interfaces would be impossible. Stanovich (2000), Shatil and Share (2003) and Leiken, Share and Schwartz (2005) are perhaps the most explicit proposals in the field of reading research along these lines and serve as another important starting point in our discussion. Contrary to strong constructivist assumptions, demarcating processing at the word identification level from that at the text comprehension level is important for understanding how different networks of components come together in each case. For word recognition, encapsulated processing applies in a manner that is different from how higher-order strategies are brought to bear on the tasks of sentence and text comprehension, according to this view. Thus, reading ability is seen as internally complex, structurally diverse.

The long-standing issue of “phonological mediation” in lexical access, following Stanovich’s approach, would now seem to be more tractable. Especially in regard to our topic, cross-language and cross-writing system literacy, an analytical method of studying how phonological structure interfaces with the other sub-components of word recognition skill should uncover interesting patterns of variation from one orthography to another. At the same time, this approach should help us refine exactly what a Universal Phonological Principle (Perfetti & Liu, 2005) leads us to claim about what is invariant in how children learn to read. The specific claims about what is invariant and what isn’t then could be tested in both first and second language literacy.

The overall idea of the UPP is that learning to read consists of learning how a writing system encodes the spoken language system associated with it, as opposed to encoding meaning directly. “Word reading activates phonology at the lowest level of language allowed by the writing system.” As a corollary, the “identification-with-phonology-hypothesis places phonology as a constituent of word recognition… that the identification of a word is the retrieval of its linguistic identity (phonologically specified morpheme of word)” (Perfetti & Liu, 2005: 195). Thus, the study of how phonological structure interfaces with other sub-components of word recognition should be relevant to the claims of the UPP.

For all of the above, so far, Jackendoff (2002) has offered a usable model of language processing that appears to be particularly apt for the study of complex abilities. Its main virtue for this purpose is that it embraces interactivity (recalling that this is not the same as holistic integration) in both feed-forward and feedback. Modules each maintain their own domain-specific computational properties, specialized for phonological structure, syntactic structure, and so forth, with interface components doing the hard work of coordinating in performance the corresponding networks and connections. Performance of language tasks is effected so rapidly and automatically that complex literacy skills, for example, appear superficially to behave like modules themselves.

The problem of the role of context is related to Stanovich’s (2000) discussion of degrees of autonomy and encapsulation, how the different components of reading ability comply with their specialized functions. The findings from research on the differential effects of context and background knowledge are proposed to be consistent with the general notion of domain
specificity: for the efficient processing of information provided to the lower-level decoding modules, significant advantages accrue if they are insulated to a certain degree, free from having to compute inputs from all possible sources. Context and background knowledge do not “pre-select” the output of word recognition, but rather intervene primarily in the subsequent stages of the reading process (Perfetti, 1992; Stanovich & Stanovich, 1999). In short, modularity in no way excludes top-down interaction; rather, what it proposes is that input computations do not have unlimited access to the higher-level domains (Coltheart, 1999).

A different dimension of componentiality poses a problem for researchers when there are two languages. Accumulating evidence now supports the hypothesis that separate linguistic systems, which correspond to knowledge of specific languages, begin to differentiate in bilingual children, early in development (Kovelman, Baker & Petitto, 2008). Conceivably, a similar differentiation would unfold in second language learning (Paradis, 2004; Genesee, 2002). But as an autonomous cognitive domain, separate from the grammatical components, conceptual structure appears not to undergo this kind of division, remaining independent and “shared in common.” See Francis (2004, 2008) for alternatives for modeling this relationship in bilingualism based on Jackendoff’s Tripartite Parallel Architecture, and Cheung and Lin (2005) for a proposal along the same lines.

The problem is now applied to the development of literacy in two languages, or in a second language. Which aspects of reading ability turn out to be language-specific and which form part of a Common Underlying Proficiency (Cummins, 2000)? For example, phonological competence per se would differentiate between L1 and L2 (or L_a and L_b in simultaneous bilingualism); but are some aspects of phonological processing skill and explicit meta-level phonological knowledge “available” from a common store, so to speak, in bilingual/cross-language situations?

In regard to question 3, strong top-down holistic theories of reading also tend to conceptually integrate language acquisition and literacy learning, rejecting the dichotomy between speech as a product of biological evolution and literacy as a cultural innovation. Opposing views point to research findings demonstrating that key aspects of metalinguistic awareness, which correlate with literacy, do not emerge “naturally” and spontaneously, as all aspects of speech-related phonological competence, for example, do. Liberman (1999, p. 108), in arguing against the idea that learning to read should proceed just as children “learn” to understand language, made the observation that:

Phonemic awareness does not result from learning to speak because the primary representations of the phonetic module are already perfectly suited for the other processes of the language specialization. These primary phonetic representations do not get attention because they do not need it, and they do not need it because…they need not be converted into something other than what they already are.
According to this view, core linguistic competence (and the ability to use language in face-to-face conversational discourse) is biologically primary, part of our “species endowment” and a constitutive primitive of human cognition. The ability to read and write is a secondary achievement, historically (the mastery of a “technology of language” to paraphrase Walter Ong, 1982). It is also a secondary achievement developmentally in children (Perfetti, 2003; Pesetsky, 1996).

A contrasting assessment of the merits of modular and connectionist approaches (question 4) is far beyond the scope of this review. However, one observation needs to be made before proceeding to the next section. The research on reading development has drawn the lines of debate in a way that perhaps is surprising. For example, on the question of the role that phonology plays in word identification and the learning of grapheme-phoneme correspondences, connectionist and modularity-oriented researchers tend to come to the same conclusion: that strong top-down constructivist theories, that minimize the importance of orthography-phonology mapping, are seriously flawed (Rayner, Foorman, Perfetti, Pesetsky & Seidenberg, 2001; Seidenberg, 2005). This coincidence between seemingly very divergent paradigms should motivate reflection on which aspects of each approach lead to this kind of convergence on specific applied problems. One possibility to explore might be related to the distinction, proposed above, between interaction and integration. A number of investigators have argued that there is no compelling reason for viewing connectionist methods and modularity as counterposed in all respects (Norris, 1990; Pinker, 1999). In the basic research on word recognition, for example, it is noteworthy that despite highly divergent approaches (Plaut, McClelland, Seidenberg & Patterson, 1996; Coltheart, Rastle, Conrad, Langdon & Ziegler, 2001), conclusions and implications for literacy learning, and teaching, diverge much less than we might expect.

**Linguistic and non-linguistic components in bilingual reading**

In their review of the research on reading in a second language, Grabe & Stoller (2002) highlight the persistent themes around which, it appears, a measure of consensus has begun to coalesce. Pertinent to our discussion of the analysis of reading ability into its knowledge and processing components are the following:

1. the distinction between tacit knowledge of language and metalinguistic knowledge,
2. how knowledge of L1 orthography influences L2 decoding (e.g. in transfer and interference), and
3. findings related to the Language Threshold Hypothesis (LTH).

An interesting aspect of the distinction in (1) - “implicit-explicit” in the previous section - is that, on the one hand, second language readers will often have an advantage in the development of higher-order literacy-related metalinguistic knowledge, in particular older learners already literate in their L1. In contrast, they will tend to be limited, to one degree or another, in their ability to fully exploit core grammatical knowledge in reading, for which the L1 reader has tacit knowledge, available on-line rapidly and automatically (Cummins, 2000; Grabe & Stoller, 2002; Walter, 2004). Depending on the circumstances, L1 and
L2 reading each might recruit resources in different proportions from the same knowledge structures.

The research on L1 writing system knowledge and L1 linguistic knowledge influences on second language reading, related to (2), has contributed to the discussion of broad theoretical problems concerning the mental architecture of bilingual competence (Wang, Koda & Perfetti, 2003). How should the mutual influences between the L1 and L2 linguistic systems be modeled? This kind of interface (or interaction) would be different from the access that L2 and bilingual readers have to a common store of higher-order, metalinguistic and discourse-level knowledge structures. Included among the latter might also be the reading-specific abilities, in particular in the domain of phonological processing and phonological awareness, which research suggests are readily accessible (termed “transferable,” in the current literature) in performance in either L1 or L2 reading (Bialystok, McBride-Chang & Luk, 2005; Chen, Ku, Koyama, Anderson & Li, 2008). Interestingly, this would be the case independent of which was the language of initial literacy learning (Geva & Wang, 2001; Hu, 2003; Mumtaz & Humphreys, 2001). The same representation and access questions arise in regard to the deployment of other non-linguistic knowledge structures. For example, research on second language reading comprehension (Morrison, 2004; Walter, 2004) has largely directed its attention to what we could characterize as abilities that are clearly not language-specific, and the extent to which these abilities can be called upon in L1 and L2 (the “transfer” metaphor commonly applied to this type of interface as well). The influence of L1 orthographic knowledge on L2 reading should now be understood better: within the framework of an analysis that conceives of components of reading ability, which component structures correspond to different kinds of knowledge and processing mechanism?

The Language Threshold Hypothesis addresses the relative weight that should be assigned to L2 linguistic competence, on the one hand, and to general reading ability and general information processing skills, on the other. The latter would have been acquired previously from L1 literacy, or from some other source in the case of non-literate L2 literacy learners. Perhaps, a better way to understand the LTH is that assuming other factors (relevant text-related background knowledge, IQ, etc.) to be constant, L2 grammatical competence and L2 lexical knowledge, up to a certain threshold, represent significant factors in reading performance. Related to this idea, Verhoeven’s (2000) “componential analysis” of L2 reading emphasizes the “dual-task,” that places greater or lesser strains on all aspects of reading, with the possible exception of the lowest levels of decoding. Even here, the critical experiments may not have yet been carried out: for example, word recognition tasks that compare the performance of native speakers with that of L2 learners below a sufficiently low “language threshold.”

In another componential analysis of L2 reading that sought to test the LTH, Leikin et al. (2005) sharpen the focus on the critical dimension of phonological processing abilities. Methodologically, the study stands out because it controlled for prior L1 literacy attainment. Investigators addressed a factor related to the idea of linguistic thresholds that might account for difficulties among unsuccessful readers: level of L2 (Hebrew) linguistic competence.
Measures of linguistic competence in Hebrew correlated with L2 reading ability. Second language grammatical knowledge was predictive, in particular, of reading comprehension. In a separate analysis of the results from a series of phonological tasks, the authors underscore the important distinction, often left ambiguous in the research literature, between two separate categories of ability, what we could term “primary” and “secondary.” Assessment procedures and interpretation of results must distinguish between: (1) core phonological competence and the “basic” phonological processing modules that deploy this competence in speech and auditory reception (“primary” ability), and (2) meta-level awareness of phonology and literacy-related processing skills (“secondary”). The investigators emphasize the need to attend to the distinction between linguistic and metalinguistic knowledge and their respective processing mechanisms and interfaces. Before moving on to the comparison between alphabetic and non-alphabetic literacy, we could say that the emphasis in the LTH on the importance of the components of linguistic competence in reading is in line with the emphasis in the UPP on one of these components in particular, phonology.

**Comparative writing system research**

The second kind of cross-language comparison that was alluded to at the beginning of this paper is the one that doesn't necessarily involve the study of bilingual readers. It is the alphabetic/non-alphabetic dimension that has attracted growing interest among researchers, for one reason because it represents the maximum possible contrast among the world's modern languages and their corresponding orthographies. For now, we will restrict the discussion to the psycholinguistics of Chinese writing and literacy development.

*Access to phonology when phonemes are not represented in the orthography*

From a number of points of view, it would seem that the literacy development of children learning the morpho-syllabic/logographic Chinese writing system could be taken as the most stringent test for the Universal Phonological Principle. In fact, support exists for the view that non-alphabetic systems implement a direct visual-to-meaning connection that circumvents phonology, the “phonology by-pass” or “direct access” hypothesis (Chen, 1996). A strong version of this proposal would be that the only necessary route from orthography (O) to semantics (S) in silent reading is the one that is not mediated by phonology (P): O → S is the dominant route. The strongest claim, a second position, would be that there is no O → P → S route at all. Phonology might be activated, but as a non-essential by-product, for the purpose, perhaps, of holding words in working memory while contextual information is processed. Taft and van Graan (1998) seem to apply a version of this model to reading in all writing systems. A third position, known as the “parallel-access” hypothesis, favors a two-pathway model: one route directly from orthography to semantics with the other mediated by phonology (Xu, Pollatsek & Potter, 1999). Simplification aside, these models, in addition to a number of subtle distinctions and inevitable qualifications, are currently being evaluated by investigators to determine what is precisely the role of phonology in reading Chinese. Componential analyses of reading ability, it will
be proposed, might be especially useful in sorting out the key research questions. One possible avenue of analysis, alluded to in the previous section, could attempt to differentiate between phonological competence and the interfaces and processing mechanisms associated with all speech and auditory reception on the one hand, and meta-level knowledge and ability, on the other.

There is another way of thinking about these three hypotheses. If Chinese presents the critical test of the UPP, it also provides for a unique opportunity for garnering support for the O \( \rightarrow \) S direct-access hypothesis (Chen, 1996), or at least a weaker version of it: accepting (provisionally) phonological activation in alphabetic reading while positing exceptionality for the processing of logographic orthographies. Assuming that this way of framing the discussion is properly conceived, what follows is a tentative proposal for why the question of phonological activation is important: Finding the direct-access hypothesis to be correct and finding that all variants of the UPP to be false, for Chinese, would contradict modular approaches to reading. Far from being a very convincing argument against direct-access, it would stand as strong evidence against modularity, as was suggested earlier. This approach to the discussion may turn out to be unsatisfactory because research findings often place different emphases and weights on the factor of phonological activation, qualifying claims that at first seemed to represent a clear counterposition. But for now, it serves us as a way of getting one view of the bigger picture.

Beginning with the direct-access hypothesis, the wide disparity between Chinese and even the “deepest” alphabet regarding the reliability of phonological information in the orthography poses, by itself, the question of a phonology bypass. On this point, there should be a relatively broad agreement among supporters of all positions in the debate (for a possible dissenting view, see DeFrancis, 1989). It has been estimated that only about 23% of semantic-phonetic compounds that Chinese-speaking children are taught during initial literacy instruction are fully regular (42% are semiregular), semantic-phonetic compounds themselves counting for less than 75% of the characters that primary school age children typically learn in school (Shu, Chen, Anderson, Wu & Xuan, 2002). In addition, a significant number of integrated characters, corresponding to free morphemes, and semantic radicals of compound characters, give the reader a graphically distinctive categorical cue to meaning, some of them containing relatively transparent iconic features (e.g. the adpositions 下 [below], 上 [above], 中 [middle/inside]). The shared semantic radical, 口 [mouth], in 吟 [chant], 吃 [eat], 吐 [vomit], and 可口 [palatable], provides the reader with semantic information, directly.\(^3\) The first two characters in the phonetic rendering of Coca-Cola, 可口可楽 [ke\(^3\)-kou\(^3\)-ke\(^3\)-le\(^4\)], correspond, opportunely, to “palatable.” Parenthetically, second language learners of Chinese have been know to make use of these visible and explicit links to meaning in the orthography (when they can) to help bootstrap themselves into a rudimentary interlanguage grammar (De Courcy, 2002, pp. 108-113, 126-127).

Chen (1996) summarizes evidence that favors direct access: In studies of word recognition, the processing of characters and alphabetic words differs
cognitively in important ways. In naming aloud and lexical decision tasks, strategies appear to differ between Chinese characters and alphabetic words. While naming is normally faster than lexical decision, and the word frequency effect is greater for lexical decision, in reading alphabetic words, the opposite result is obtained with Chinese characters - lexical decision is usually faster and frequency effects are greater in the naming task. In an experiment that required Chinese/English bilinguals and English monolinguals to identify mathematical logographs and corresponding words, logographs and characters were named equally fast by bilinguals, while word-naming was faster for the alphabetic-reading monolinguals. Interestingly, when the bilinguals attempted the task in their L2 (English), performance resembled that of the native-speaker monolinguals. Together with findings that show pre-lexical phonological activation to be weaker in Chinese than in English, and other evidence indicating a lack of homophonic effects in semantic categorization, an important contrast between alphabetic and non-alphabetic processing appears to have been demonstrated (Chen, 1996, pp. 51-56). This contrast appears to suggest that the O → S route predominates.

For some versions of the direct-access hypothesis, what is called into question is not that phonological structures are automatically activated in silent reading. Even an early activation of phonology in Chinese word identification might be possible. Rather, the claim would be that mediation of phonology does not strongly constrain access to meaning. For example, studies have shown significant effects of phonology for only a minority of Chinese characters in which the pronunciation of the compound character is the same as the pronunciation of its phonetic component (Zhou & Marslen-Wilson, 1999). According to this view, the O → P → S (mediated) route: “cannot be the predominant or default mechanism for linking orthographic form to lexical semantic representation. There is little doubt that phonology is activated early and obligatorily in reading Chinese characters...[and] that under specific circumstances phonological factors can drive semantic activation for skilled readers. It is, however, much more plausible to view phonological and orthographic factors as functioning in an interactive framework...” (p. 598).

Zhou and Marslen-Wilson's position appears to be close to the two-pathway parallel-access view. According to Xu et al. (1999), a model in which orthography accesses phonological and semantic codes in parallel has the virtue of not assuming different processing mechanisms for Chinese and alphabetic systems. Findings that indicate a divergence between word reading processes in the two systems can be explained simply by the relative speed of the route in which phonology is more easily activated by an alphabet versus the route to meaning in Chinese in which the O → P → S route would end up being slower than the direct O → S pathway. Here again, the same phonology-semantics connections of speech comprehension come on-line automatically, and are not by-passed, strictly speaking. An interactive network for literacy would consist of three “nodes,” one for orthography, one for phonology, and one for semantics. During reading, all three would be in continuous interaction (p. 853).
Lexical entries in a mental lexicon

As we can now well appreciate, adjudicating among the various claims regarding the role of phonology in Chinese reading will depend on what comes to be considered as the central distinguishing feature(s) of each competing hypothesis. However, which features are in competition is not entirely obvious either; see Guo, Peng & Liu (2005), Tsai, Lee, Tzeng, Hung and Yen (2004), Tzeng, Zhong, Hung and Lee (1995), and Zhou and Marslen-Wilson (2000) for discussion and relevant research findings. At this point, perhaps, only the confrontation between the strongest version of direct-access (Taft & van Graan, 1998) and models roughly convergent with the UPP will allow for a first approximation of a definitive test. This being a good place to begin, Packard (2000) seems to dispense with the notion of “routes” (O \(\rightarrow\) S, O \(\rightarrow\) P \(\rightarrow\) S), suggesting rather an alternative based on a model of the mental lexicon with each lexical entry composed of sub-structures. The lexical entry is a “relation” among: phonology, syntactic form class and subcategory, morphology, semantics, etc., interfaced with an orthographic representation for literate speakers. In the end, these may not be the actual sub-structures; but the basic idea would be the same for explaining how Chinese characters activate phonology and meaning.

A brief overview of Packard's model of how characters access the mental lexicon will help frame our concluding assessment of the Universal Phonological Principle. The model assumes that Chinese orthography is completely dependent upon the natural speech lexicon. The linguistic sub-systems that are accessed in reading are the same ones that all speakers of the language possess, including non-literate. It would be less plausible that literacy development recreates a parallel, duplicate, competence for processing written language. What literate speakers of Chinese are typically able to count on is a more advanced explicit awareness of morphemes and words, that might help them with certain analytic tasks and in reading itself, especially in the case of difficult texts. But, the development of this metalinguistic knowledge remains separate from, and leaves intact, the core components of native-speaker linguistic competence, including all interface modules. For example, in bilinguals, script differences (e.g. in English-Japanese bilinguals) apparently do not block cross-language influences that are typically expected in “same-script” bilinguals (English-Spanish) (Hoshino & Kroll, 2008).

Recall the distinction made by Liberman (1999) in the first section between primary linguistic knowledge and secondary, literacy-related, meta-level knowledge. Following this logic, the grammatical subsystems of the lexical entry are tightly interlinked, input of a linguistic form (orthographic in this case) triggering the “retrieval” of its core linguistic sub-structures, rapidly and automatically in skilled reading. In speech perception, access is triggered by sound, which can only activate the sub-structures of the lexical entry by way of phonology. In reading, on the other hand, lexical entries could be activated by triggering another sub-structure “first”; and as Packard (2000) pointed out, in Chinese this might even be more likely under certain word identification task conditions. Here again, we have an indication that modular approaches (such as
those of Liberman and Packard) lend support to the UPP, and are not consistent with strong versions of direct access (i.e. phonology by-pass).

Consistent with this view, studies of dyslexia, dysgraphia and aphasia have shown that the lexical sub-structures and the corresponding linguistic sub-systems and interfaces with which they are linked can suffer selective impairment. In the case of dyslexia, different networks of processing mechanism can be affected yielding a diverse array of impairments, not a single undifferentiated category of reading impairment (Chung, McBride-Chang, Wong, Cheung, Penny & Ho, 2008). And as would be expected, different patterns of disability are associated with alphabetic and non-alphabetic dyslexic reading (Ho, Chan, Lee, Tsang & Luan, 2004). Studies have revealed opposite and complementary patterns of selective impairment: fluent speech accompanied by disabled oral reading (a defective orthography-phonology link) and disrupted access to meaning in speech with spared oral reading (suggesting an intact orthography-phonology pathway and a disrupted access to semantics) (Weekes, Yin, Su & Chen, 2006; Law, Wong & Chiu, 2005). With phonology and semantics being dissociable, nothing in principle then should exclude the possibility of an initial activation of either one under experimental conditions of single character or single word identification.

In unimpaired readers, however, the interfaces among the linguistic sub-components of the lexical entry must be dedicated, fast and obligatory in their activation, revealing properties of highly encapsulated modular processing; their deployment is mandatory. These inter-connections within the natural lexical entry have evolved, hypothetically, as part of the specialized language faculty and would be highly resistant to being reconfigured in any fundamental way. Crucially, the connection to orthography (established through learning and extended practice) is not predetermined in this way. Therefore, a “direct” initial access to a word's meaning, demonstrable in experimental tasks, cannot be effected independently of the phonological structures of spoken language in the sense of by-passing them; in this sense, “recoding” is a “virtual reflex” (Packard, 2000, pp. 304-309). In actual reading of continuous text, if any cognitive domain is susceptible to disengagement it would be a higher-order level of conceptual structure related to interpretation and sentence/text comprehension, a phenomenon that metacognitively aware readers are quite familiar with (the reverse, not coincidentally, of the well-known tip-of-the-tongue state). Under normal receptive language processing, given a fragmentary structure in a particular format, the interface modules work rapidly to construct a fully specified structure (Jackendoff, 2002, p. 198), presumably in word identification just as in syntax. The lexicon as a whole, according to this view, is a substantive part of the language processor's network of interface components:

A word, by virtue of having features in each of the components of grammar, serves as part of the linkage between the multiple structures… The proper way to regard [the lexical item] is as a small-scale three way interface rule. It lists a small chunk of phonology, a small chunk of syntax, and a small chunk of semantics, and it shows how to line these chunks up
when they appear in parallel phonological, syntactic, and conceptual structures (p. 131).

**Characters and words in context**

One could perhaps question experimental methods that are restricted to studying individual word (and non-word) identification on the grounds that they lack ecological validity (Goodman, 2005). However, this would be an unfair criticism that confounds important methodological considerations of data collection, and the complementary but different requirements of reliability and interpretation (validity). For example, if consistent results demonstrate significant differences between the processing of individual words in Chinese and a given alphabetic orthography, the results stand as reported, and invite theoretically grounded explanation, speculation, and methodological critique. Competing accounts of the data inform subsequent experimental and non-experimental lines of inquiry; and if one chooses to dismiss a given method of inquiry, it is necessary to explain why it cannot yield any interpretable finding, under any circumstance.

At the same time, studies of reading that require subjects to decode connected text should complement single-character/word-in-isolation studies in all respects. Lin and Akamatsu (1997) appeared to suggest this approach in their observations of the research on eye movement patterns and other aspects of reading continuous text; see Frenck-Mestre (2005) for discussion. As expected, comparative eye movement studies reveal noteworthy differences between Chinese and English. In reading “Chinese in context,” nevertheless, phonological representations are activated and play an important role, even if we assume, for argument's sake, that activation is “post-lexical.” In this case, it is proposed, we should assume that the same processing mechanisms are brought on-line in Chinese and English. What appears to follow is the conclusion that text processing in all languages and across all writing systems cannot by-pass phonological recoding for the purpose of syntactic and semantic integration in linguistic working memory. However, this conclusion should not be taken as contradicting the evidence from numerous studies “that skilled readers of different orthographies [adapt] their processing strategies to meet the different cognitive demands posed by different orthographies” (Lin & Akamatsu, 1997, p. 378).

In their study of the development of early reading ability comparing performance in the Taiwanese Mandarin alphabet, Zhuyin fuhao, and logographic Chinese, Hu and Catts (1998) comment on the role of phonological working memory. In particular, the issue comes up in decoding less familiar words in context. Since Chinese orthography does not indicate word boundaries, phonological patterns need to be maintained in working memory until word boundaries are identified. Less skilled beginning readers, it has been proposed, have difficulty in segmenting effectively between juxtaposed characters, distinguishing in on-line processing of text between free and bound characters and parsing grammatical constituents correctly that contain two-character and multi-character compound words. In this way, phonology is activated and plays an important role in decoding.
Independently of the mixed results from single-character/word-in-isolation studies on whether phonology is “prelexical” or “postlexical,” in the reading of connected text, what is the effect of an ongoing and overriding engagement of phonological working memory? If the above observations are correct, can we even plausibly pose the possibility of a direct access to meaning without phonological mediation in word identification in text processing? These two questions are indirectly suggested in the discussion of methodological problems in single-character/word-in-isolation studies. Chen (1996) contrasts two kinds of approach: (1) tasks are presented to the subject in which a character's name is required or can be helpful, and (2) situations in which the meaning of a character is to be consulted for a correct response with phonological information unnecessary, as in the semantic-decision task. In (2), if no effect of phonological interference is in evidence then one explanation is that phonological activation does not occur routinely or necessarily as a part of character recognition (Chen, 1996, p. 53). But, in reference specifically to the second scenario, if in context the phonological component of linguistic working memory comes on-line mandatorily (even if access to phonology on the first character of a text passage could be “post-lexical”) direct, unmediated, access to meaning may be impossible. Entirely speculative, this possibility may be worth exploring in future investigations. But again, in relation to (2), if significant differences between Chinese and English, for example, continue to be evidenced, this finding requires further serious discussion, even if it turns out that “no effect of phonological interference” is only found in single-character/word-in-isolation tasks.

*Phonology as a constituent of word identification*

The Universal Phonological Principle was characterized as a “constituency hypothesis” by Guo et al. (2005). A more complete summary of how the UPP should or might apply to Chinese now merits our attention. Perfetti (2003) began by pointing out the incomplete and thus misleading conception of the character as morphemic. Rather, “morphosyllabic” better captures its orthographic-linguistic properties: “corresponding not to an abstract formless piece of meaning, but usually to a spoken Chinese syllable that is also a morpheme” (p. 7). On this point, a morphosyllabic writing system is not exceptional; it also realizes a mapping between orthography and linguistic patterns of a language. Far from a quibble about terminology, this assumption about how a “true writing system” functions (DeFrancis, 1996) points us toward an important ongoing discussion in cognitive science: assuming that the semantic domains of conceptual structure (Jackendoff’s term, 2002), are autonomous from the linguistic structures (morphology, syntax, phonology), how do meaning and language, now defined narrowly, interface in language use (such as in reading)? The more fundamental debate revolves around the proposals of both meaning-linguistic structure differentiation and the distinction between knowledge of language and language use (Newmeyer, 1998), both of which also have been widely challenged.

Assuming for now that in silent reading the activation of phonology occurs in both alphabetic and non-alphabetic systems, it would be incorrect to claim no difference in the execution of the decoding processing mechanism. As
we could see from the brief review of the research evidence in the previous section, on this point there seems to be wide agreement. As Perfetti describes it, an alphabetic system allows activation to occur sublexically based on letter recognition, building up rapidly; and word identification does not have to wait for complete processing of all lower-level units. In silent reading of Chinese, in contrast, activation of phonology awaits a threshold-level of character recognition (see Note 3). Summing up, we can consider the following contrast: alphabetic-cascade style, logographic/morphosyllabic-threshold style. Thus, the basic idea of phonology being a constituent of word identification is that the moment of orthographic recognition is the moment of access to phonology, constituting together with the lexical entry's semantic features a “three-constituent word identity” (Perfetti, 1999, p. 177). This would stand in contrast to phonology as a by-product of identification, “post-lexically.” In English, the graphemes that trigger phonological processing correspond to phonemes; in Chinese they correspond to spoken syllables that happen also to be morphemes (Perfetti & Liu, 2005, p. 195).

The intention in all of the preceding discussion is not to argue that the Universal Phonological Principle has already garnered decisive support as the best theory for Chinese literacy (if we recall, this being the most demanding test), but rather that it is broadly compatible with modular approaches to the study of language ability. If it turned out to be substantively correct, then this would favor, according to the argument, at least some versions of the modularity thesis. Another way to appreciate the argument might be to ask: why, typically, do strong top-down holistic theories of reading tend to favor models in which semantic/pragmatic context and background knowledge penetrate decoding extensively; and why do they tend to minimize the importance of bottom-up mechanisms, in regard to the development of phonological/orthographic skills in particular? Also strongly anti-modular, they appear to coalesce around a kind of “direct-access” model, implying a direct interface (or degrees of “directness” depending on how strongly holistic one's position is) between orthography and semantics. Antithetical (to holism) modular models of reading, that for example are “subskill-oriented,” pedagogically, appear to be strongly bottom-up. This is because, following the assumptions of the modularity approach, certain components of language processing in reading are engaged in a mandatory way, not subject to by-pass or deflection by “central” top-down influences.

A pivotal concept in understanding these differences is that of mental representation, that actual neurological structures subserve competencies and processors. What does it mean when we refer to cognitive components of reading, for example, in discussing the research on feedback influences (from higher levels) on lower level representations? How is it that in some cases the “feedback loops” are effective, in some cases less so, and in others ineffective because feedforward (relatively encapsulated) processes predominate (Perfetti, 1999, p. 171)? Along with the conception of the lexicon as a mental representation of a special kind (a part of the domain of interfaces), are these components intrinsic and constitutive of a “permanent” language faculty, or should they be considered more like emergent patterns of activation (not “permanent”? The following
concluding section on developmental aspects of Chinese literacy might provide another point of view on the role of phonology in reading.

**Conclusion: Literacy-related skills and awareness**

To wrap up this review of the research we turn our attention to a practical issue: that of phonological skills (and explicit knowledge about phonology) in learning how to read. One could accept the phonology-as-constituent-of-word-identification hypothesis (the UPP), and be neutral regarding the role that phonological awareness plays in literacy learning, or take one position or another. The componential approach that we have been considering as an overall analytic framework might help to explain why. To start with, the difference between the two questions is mainly one between a representation problem and a learnability problem. The distinction between explicit knowledge/metalinguistic awareness and implicit competence (question 1 on the list of the four modularity-related research questions from the first section) also allows us to separate the two issues: meta-level abilities associated with phonology are not relevant, or not very much, in evaluating the claims of the UPP. Evidence either for or against the claim can be drawn from the study of the activation of the encapsulated phonological module and performance in a “closed” sub-system. In contrast, the study of the role of phonological awareness in literacy learning focuses on the development of metalinguistic knowledge in a domain that is “open” and subject to deliberate reflection. Fortunately, unlike the contention among the different claims regarding phonological mediation in word identification, the lines are drawn much more clearly on the issue if how explicit knowledge develops. We can outline them concisely as follows.

There are different viewpoints on the importance of phonological awareness skills in child literacy learning. But there is one result around which some measure of agreement exists. It appears that learners decompose compound characters, and that they make use of and are sensitive to different degrees of partial information found in the phonetic component. There may be a stage in reading development in which an analytic stance toward the phonetic component of compound characters facilitates children's mastery of the orthography (Shu, 2003; Tzeng, 2002). It has also been suggested that Chinese dyslexic readers have difficulty applying phonological processing strategies, and that direct teaching of script-sound regularities improves reading performance in these cases (Ho & Ma, 1999). All of this might be true even if it turns out that phonological awareness and phonological processing skills, overall, play a role in Chinese literacy that is significantly different from that in alphabetic systems (Ko & Tzeng, 2000). Morphological awareness, for example, might play a much more prominent role in literacy learning (Li, Anderson, Nagy & Zhang, 2002; Shu, McBride-Chang, Wu & Liu, 2006).

One interpretation of this finding is that the development of this kind of phonological awareness gets its boost mainly from literacy learning itself, or that there is a complex reciprocal relationship subsequent to initial experience with literacy and literacy-related language use in general. This argument has been well developed in the research literature on alphabetic literacy learning over the years;
see Castles and Coltheart (2004) and Ziegler and Goswami (2005) for a comprehensive review. Long ago, Perfetti (1992) observed that: “Downgrading phonological awareness from causal status to reciprocal status does not diminish its importance for reading. Indeed, it allows it to be seen as a central component of reading instead of as a prerequisite” (p. 166).

Among researchers who question the hypothesis that phonological awareness, per se, predicts successful literacy learning, we might be able to discern two approaches: (1) that the hypothesis should be disfavored in general, and (2) that it can be discarded in the case of non-alphabetic literacy learning (in regard to Chinese, we should assume that the consideration of the factor of phonological awareness is restricted to the syllable level, not to phonemic awareness). For now, we will set aside the first approach and consider the possibility that in Chinese the relationship between reading development and phonological awareness is secondary or even marginal. This was the conclusion of Tan, Spinks, Eden, Perfetti and Siok (2005) in a study that attempted to analyze the “component skills” that are relevant to the complex ability of character recognition. Recalling that children must grapple with “system-level” complications (both linguistic and orthographic) in Chinese that make mapping between characters and linguistic and meaning structures significantly more complex than in any alphabetic system, the first assumption might be that the learning task is unlikely to be supported by phonological awareness in a major way. The authors propose that the formation of an “integrated reading circuit that links orthography, meaning and pronunciation” (Tan et al., p. 8781) occurs at the visual-orthographic level facilitating children's awareness of the character’s internal structure. This “orthographic awareness” appears to depend on the establishment of long-term motor memories of characters through extensive practice in copying and writing. Naming speed and measures of accuracy in copying characters were found to be strongly related to reading ability; the relationship with measures of phonological awareness (oddity test and syllable deletion) was minor. Addressing the developmental/learning issue, these findings confirm the results of a number of other studies, suggesting an important difference on this point between alphabetic and non-alphabetic literacy development (Taylor, 2002).

A number of studies make a strong case for an alternative view. Starting from the idea that children might implicitly “analyze” patterns in the internal structure of characters and make use of both the phonetic in fully regular compounds and partial information elsewhere, researchers have proposed a greater role for awareness of phonology (at the syllable level, not to forget). Evidence from studies of the regularity effect (more accuracy in reading regular characters), analysis of reading errors (overgeneralization) and pseudo-character naming (positively correlated with reading achievement) suggests the development of a “phonetic principle.” This aspect of phonological awareness helps children store compound characters more systematically and facilitates the learning of new characters (Chen, Shu, Wu & Anderson, 2003), for example, by using an “analogy strategy,” decoding unfamiliar characters with recourse to
knowledge of the same phonetic of a familiar character (He, Wang & Anderson, 2005).

While findings on the relationship between phonological processing, phonological awareness and Chinese reading remain mixed, the number of studies that have found a positive correlation, contrary to Tan et al. (2005), for example, is noteworthy and future work in this important domain of literacy learning will need to reconcile the differences. Memory tasks involving sets of nonwords, sound categorization (which word out of a group is different) (Hu & Catts, 1998) and syllable deletion (Chow, McBride-Chang & Burgess, 2005; McBride-Chang & Ho, 2000) have been shown to be significant predictors of reading ability in young children. Consistent with other biliiteracy studies, the Chow et al. study found that performance on syllable detection in L1 (Cantonese) also predicted L2 reading (in English). Siok and Fletcher (2001) found that onset-rime awareness was positively related to reading ability at the higher grades (visual skills at the lower grades), but phonemic awareness, as we would predict, was not. As was noted earlier, there is a general recognition of the reciprocity between literacy learning and phonological awareness. However, as a component of early reading development (even if it cannot be shown to be a prerequisite) the participation of phonological awareness and the importance of phonological processing are not minor or secondary for the above mentioned researchers who appear to have drawn different conclusions than the ones in the Tan et al. (2005) study. In other words, the importance of phonological awareness and phonological processing is not specific to alphabetic literacy learning (Chan & Siegel, 2001). In conclusion, the most promising first step toward reconciling the different positions might be to systematically confront the opposing findings from research on the development of reading ability, including phonological awareness (a learnability problem) among studies that all assume one version or another of the Universal Phonological Principle (mainly a representation problem).

Notes
1. Strong views of modularity tend to disfavor extensive inter-modal or cross-domain interaction (Fodor, 1998). Weaker versions allow for a conception of modules (knowledge structures and processing mechanisms) as “encapsulated” to a lesser extent (Francis, 2008; Marcus, 2006), and for cognitive components to show degrees of “domain-specificity” or “domain-generality.” Thus for the latter, the degree of interactivity among domains is more of an open question, yet to be determined by empirical research. What weaker versions would reject is that interactivity is unconstrained and that mental architecture is entirely homogenous and holistically structured, the key distinction then being between interaction (among components/modules) and non-componential integration (no domain-specific cognitive structures). Jackendoff (2002) discusses at length the different versions of modularity in regard to these dimensions.

2. The Language Threshold Hypothesis and the "dual-task of L2 literacy" concept are at odds with proposals that minimize the factor of linguistic knowledge in learning how to read in a second language. Poorly designed studies aside, it seems
that much of the counter-evidence is based on evaluations of L2 reading ability in which subjects' L2 grammatical competence has already advanced beyond the required threshold for the specific text material or assessment task under consideration. In many studies, subjects' level of L2 grammatical competence is simply left unspecified. More controlled studies, in the coming years, should begin to converge on what has been taken as a working assumption by most researchers in the field. That even though we should recognize the effects of a number of interacting external social factors, there exists for *pre-literate monolingual L1-speaking child learners*, in addition to pre-literate L2-beginners, a Fundamental L1-L2 Literacy Difference, a position that we have argued in favor of from our work on second language and bilingual literacy in Central and Northern Mexico (Francis & Paciotto, 2004; Hamel & Francis, 2006). See the section on "Language of instruction" (D. Francis et al., 2006), in the *Report of the National Literacy Panel on Language-Minority Children and Youth* and related sections that are consistent with the Language Threshold Hypothesis and the Fundamental L1-L2 Literacy Difference, and Guglielmi (2008) and Restrepo and Gray (2007) for findings that support the facilitative effect of L1 literacy-related proficiency on L2 literacy attainment.

3. Compound characters are composed of two subcomponents (or radicals), each one usually an independent integrated character by itself. Usually they consist of a semantic radical, that often provides a categorical clue for meaning, and a phonetic radical (or phonetic) that can give or suggest pronunciation (see examples for “chant,” “eat,” and “vomit” – semantic radical on the left, phonetic on the right). Phonetics can also be “bound,” not appearing as independent characters. Some compound characters, on the other hand, combine two semantic components to suggest a new meaning. Integrated characters are not composed of separate phonetic and semantic components. They consist of stroke patterns that form an integrated unit that is inseparable. In the examples given, “below,” “above” and “middle” are independent integrated characters. The great majority of compound characters are irregular or semi-regular, neither radical providing a completely reliable clue for meaning or pronunciation. Perhaps the feature of the Chinese writing system that distinguishes it most from alphabetic systems is that there are no component parts of a character that encode individual phonemes. Also taking into account the overall inconsistency of the phonetic radical, the correct pronunciation of characters requires identification of the entire character, unlike in alphabetic writing where a reader can reliably “sound out” unknown words and provide a more or less correct pronunciation for them. (Chen, Shu, Wu, & Anderson, 2003; Perfetti, Liu & Tan, 2005; Weekes et al. 2006; Xu et al. 1999).

**References**


**Address for correspondence:**
Norbert Francis  
College of Education  
Northern Arizona University  
PO Box 5774  
Flagstaff, Arizona 86011, USA  
norbert.francis@nau.edu  
http://jan.ucc.nau.edu/nf4/