中央研究院歷史語言研究所集刊
第六十五本，第一分
出版日期：民國八十三年三月

The Structure of Orthographic Representations of Chinese Characters: from the Perspective of the Cognitive Neuropsychological Approach

Sam-po Law

In this paper, the view that Chinese characters do not carry phonological information and the view that they are pictographs are argued to be incorrect. On the contrary, Chinese orthographs have considerable phonetic information, similar to alphabetic writing systems, and they have internal structure. Evidence for this claim is drawn from findings concerning the reading performances of normal Chinese subjects and of a brain-injured patient, and the writing performance of Chinese dysgraphic patients. In particular, normal subjects were found to name phonetic compounds whose phonetic component has a high consistency value faster than those containing a phonetic component low in consistency value. This effect of consistency was also observed in the reading of pseudo-characters made up of a signific and a phonetic (Fang, Horng, and Tzeng 1986; Hue and Erickson 1988; Lian 1985). The non-phonetic view regarding the Chinese writing system contends that Chinese lacks a non-lexical route of reading; it thus implies that character reading is achieved post-lexically. However, this is seriously challenged by a report about a Chinese brain-damaged patient who could read invented as well as real characters effortlessly, despite little comprehension of the characters (Tzeng, Hung, Lu, Chen, and Hu 1992). Writing errors from several dysgraphic patients of a study investigating the writing processes in Chinese found that the signific and the phonetic may be substituted, deleted, or inserted. These errors indicate that the signific and the phonetic must correspond to units or constituents at some level of cognitive processing, suggesting that the orthographic representations of phonetic compounds contain information about the content and the position of their signific and phonetic component. The use of evidence from neuropathology in this paper illustrates how data from brain-damaged individuals may inform and constrain models of the cognitive system, as practiced by cognitive neuropsychologists.

Introduction

The writing system in Chinese has been treated by some as pictographic in nature [1] ; it is also considered a deep orthography, in the sense that the mapping between orthographs and sounds is very opaque. In this paper, I would like to argue that these views are incorrect. Contrary to these misconceptions, it will be shown that Chinese characters carry considerable phonetic information and have internal structure. This claim is justified by the reading performances of normal subjects as well as patients with acquired language disorders on naming single Chinese characters, and by a proposal about the structure of orthographic representations of phonetic compounds based on writing errors from Chinese dysgraphic patients. Since the use of data from brain-damaged patients to make inferences about the normal cognitive system stems from the cognitive neuropsychological approach, I will first describe the main objectives, basic assumptions, and methodological issues in cognitive neuropsychology as have been discussed in the literature, before reviewing research on reading of Chinese characters and presenting the proposal regarding the orthographic representations of phonetic compounds, which are characters composed of a signific component giving a clue to the meaning of the character and a phonetic component providing some cue to the sound of the character.

---

1 A recent example can be found in Hu, Qiou, and Zhong (1990).

The organization of the paper is as follows. Section 1 discusses very briefly what cognitive neuropsychology as a discipline is about, in terms of its goals, underlying assumptions, and methods. Section 2 describes results of recent research on naming of phonetic compounds. It is shown that the phonetic component plays an important role in the reading of normal subjects and aphasic patients. The results of this line of research also suggest that in the process of reading phonetic compounds, the orthographs are analyzed into components, one of which being the phonetic. This in effect argues that the representations of phonetic compounds at the input level have internal structure. Finally, in Section 3, I will present a proposal concerning the structure of output orthographic representations of phonetic compounds. The proposal is made based on writing errors produced by Chinese brain-damaged patients. The bulk of the data comes from a preliminary study on reading and writing performances of Cantonese Chinese aphasic patients recently conducted in Hong Kong. It is suggested, at a general level of description at this point, that the representations of phonetic compounds contain information about the identity and position of occurrence of the signific and the phonetic component. Evidence for this consists of error types including substitution, omission, addition, and transposition involving the signific or the phonetic. This analysis thus argues that output representations of phonetic compounds, similar to their counterparts at the input level, have internal structure.

1.0. What is Cognitive Neuropsychology?

This section intends to be a *general* characterization of the cognitive neuropsychological approach; many subtle and complicated issues concerning the methodology and underlying assumptions within this approach will not be discussed in detail. The content of the section is based heavily on the following sources: Badecker and Caramazza (1985), Campbell (1987), Caramazza (1984; 1986; 1988), Coltheart (1985), Coltheart, Patterson, and Marshall (1980), Ellis and Young (1988), McCloskey and Caramazza (1988), Saffran (1982), Shallice (1988).

1.1. Objects of Study and Objectives

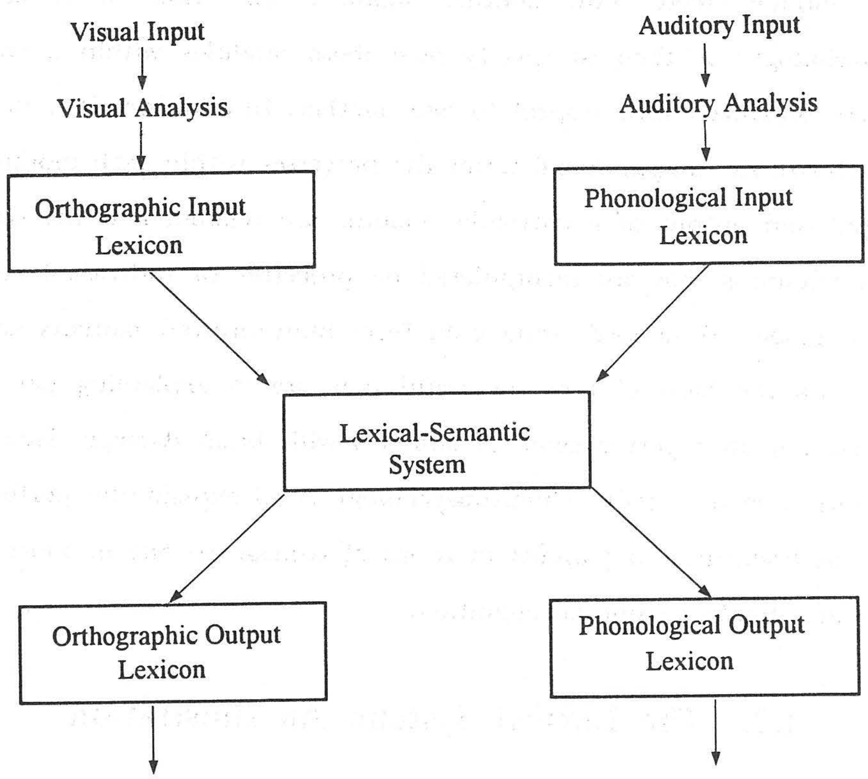
As the name itself suggests, cognitive neuropsychology is the convergence of two disciplines, i.e., cognitive psychology and neuropsychology. The former studies the mental processes that underlie normal functioning, including object recognition, speaking, reading, writing, problem solving, decision making, remembering, etc., and the latter investigates how behaviours, ranging from perception, sensation, emotion, to various mental events, are mediated by brain structures and processes. Cognitive neuropsychology then purports to research into the cognitive systems supporting various mental activities, using evidence from neuropathology. In other words, cognitive neuropsychology has the general goal of making inferences about normal cognitive functions based on patterns of impaired and preserved performance from brain-damaged patients. With an influence from the information-processing approach, this objective is taken to mean the "formulation of a functional architecture which specifies the componential structure of a cognitive system and the specification of the computational structure of the individual components of processing that comprise the functional architecture of the system" (Caramazza 1986, p. 47). Put another way, a model of the cognitive system, being modular in nature, consists of a number of components (or modules). These modules are independent in the sense that they engage in processing their own forms of information, but they may also be interrelated to the extent that the output of one module feeds into another module. The task of a cognitive neuropsychologist is then to specify how these modules within a cognitive system are organized with respect to one another. In addition, it is necessary to characterize in computational terms the processes within each module, i.e., how input and output of a particular module are related, and the structure of representations that are manipulated by processes of individual modules. Cognitive models developed using data from brain-injured patients are then evaluated on the basis of how successful they are in explaining patterns of intact and impaired performance of subjects with brain damage. Hence, the second objective of cognitive neuropsychology is to explain the performance patterns of brain-injured patients in terms of damage to one or more of the modules of a model of normal cognition.

1.2. The Lexical System: An Illustration

This section discusses some aspects of the lexical system which underlies single-word processing, in order to illustrate in more concrete terms the nature of cognitive models according to which cognitive neuropsychologists provide accounts for normal functioning and patterns of impaired and perserved performance of brain-damaged individuals.

Figure I represents a composite model concerning the recognition and production of spoken and written words.

Figure I. Schematic representation of the lexical system



The lexical system in Figure I contains a set of interconnected components (modules). This set of constituent modules indicates that the system makes two distinctions, one between input and output lexical components, another between modality-specific input or output modules. These two distinctions result in four lexical components: the phonological input lexicon which is involved in processing spoken words; the orthographic input lexicon involved in processing written lexical items; the phonological output lexicon responsible for the production of spoken words, and the orthographic output lexicon involved in the production of written words. These four modules are related through their connections to the lexical-semantic component which stores the semantic representations of lexical items.

The postulation of a lexical system with a specific set of modules and structural organization as shown in Figure I is motivated by evidence that brain-injured patients may present with selective impairment to one or more lexical components, resulting in particular performance patterns. For instance, a selective damage to the phonological input lexicon would result in an inability to comprehend most or all spoken words, although reading comprehension and speech production should be preserved. Patients who show this pattern of impaired and intact performance have been reported in Kohn and Friedman (1986) and Allport and Funnell (1981). These patients are labelled as having "word-meaning deafness". A selective damage to the lexical-semantic component would result in the production of semantic errors. According to the functional architecture of the lexical system in Figure I, the lexical-semantic component feeds into both output lexical-components; hence, the prediction is that damage to the semantic system should affect both oral and written production. Marshall and Newcombe (1973) reported two patients classified as "deep dyslexics" (a discussion of this syndrome will be provided in section 2) who made semantic errors in reading aloud, for example, reading *large* as "big", *city* as "town", *berry* as "grapes", *diamond* as "necklace", *frail* as "fragile", etc. Nolan and Caramazza (1983) found a patient who made semantic errors not only in reading aloud but also in picture naming, written tactile naming, and writing to dictation. For example, the patient wrote "trial" for *sentence*, "afraid" for *frighten*, "fight" for *argument*. On some occasions, the impairment in the lexical-semantic component is so specific that only lexical entries related to some semantic categories are affected. Warrington and McCarthy (1983) discussed a patient who presented with particular difficulty in comprehending names of *inanimate* objects. This contrasts with the patient reported in Warrington and Shallice (1984) who showed problems in processing names of *living things.* Finally, Goodman and Caramazza (1986) described a patient who had an impairment to the orthographic output lexicon with preserved functioning of other lexical components. The patient was able to spell non-words orally and in written form with an impaired capability in spelling real words. These reports serve to illustrate the use of neuropathological data in drawing inferences about normal cognition. They constitute only a small portion of a large body of literature supporting a modular view of the lexicon.

1.3. Theoretical and Methodological Assumptions

The discussion of the lexical system in the previous section has implicated the underlying assumptions one must make in order to validly relate data from brain-damaged patients with theories about the normal cognitive system. The most fundamental and least controversial is the *assumption of universality* which states that the functional architecture of the cognitive system is generally the same for normal individuals. Without this assumption, scientific investigations into normal cognition would not be possible. The second assumption, a somewhat more controversial one, is the *assumption of transparency.* As already illustrated in section 1.2. by the way in which inferences about the normal lexical system are made based on patterns of performance of patients with brain injury, the transparency assumption asserts that "the cognitive system of a brain-damaged patient is fundamentally the same as that of a normal subject except for a "local" modification of the system.." (Caramazza 1986, p. 49). Put another way, the patterns of intact and impaired performance of brain-damaged patients reflect a *discoverable and specifiable* transformation of the cognitive system. Moreover, the assumption requires that no *de novo* cognitive mechanisms are created as a result of brain damage. This is necessary because, otherwise, data from brain-damaged patients would not give us information about normal cognitive functioning.

Besides these theoretical assumptions, methodological considerations are also required to determine the proper method for cognitive neuropsychology. One may notice that cognitive neuropsychologists make hypotheses about the locus of functional lesions (which are lesions or damage to the cognitive system) on the basis of information available to them regarding the pattern of impaired and preserved capabilities of a patient. The determination of a functional lesion in any individual case can only be made *a posteriori.* In other words, the conditions under which a particular pattern of performance is resulted are not known and must be inferred. This contrasts with the situation in experimental studies using normal subjects, where the experimental conditions are always known, and are assumed to be constant across participants. An important consequence emerges from this difference, namely, the assumption of homogeneity that justifies the grouping of subjects and the averaging of performances of participants in experimental studies using normal subjects cannot be maintained in investigations involving patients. The conclusion is unavoidably that only single-patient studies can provide useful information to cognitive neuropsychologists in developing models of normal cognition (Caramazza 1986; 1988; 1992). The grouping of patients would not be justified even in cases where patients showed qualitatively equivalent patterns of performance, since one cannot guarantee that this will continue to be the case with further observation of the patients. Although the reasoning behind the conclusion is powerful and compelling, the claim that the single-case study approach is the only proper method for doing cognitive neuropsychological research is still being debated (e.g., Bates, Appelbaum, and Allard (1991); Bates, McDonald, MacWhinney, and Appelbaum (1991)).

An issue related to the justifiability of patient grouping is concerned with the value of patient classification. Objections to the use of clinical categories in cognitive neuropsychological research have been raised and discussed in detail in Caramazza (1984), Badecker and Caramazza (1985), Caramazza (1986). It was argued that the specific groupings of symptoms that define various clinical syndromes are not theoretically motivated. In addition, patients who are classified as belonging to the same syndrome category often do not show qualitatively identical patterns of performance, or do not present with the whole range of symptoms related to the syndrome in question. Furthermore, patients who are classified as having the same clinical syndrome based on the presence of some defining symptoms may well have different underlying deficits. Given that the nature of inference in the cognitive neuropsychological approach is always *a posteriori,* homogeneity across patients of a certain syndrome category in theoretically relevant respects cannot be justified. Finally, as mentioned earlier, one of the goals of cognitive neuropsychologists is to explain performance patterns of *individual* brain-damaged subjects, patient classification based on clinical syndromes provides information that is of little value to research in neurolinguistics and cognitive neuropsychology.

Summary

The cognitive neuropsychological approach has the aims of drawing conclusions about cognitive systems using evidence from neuropathology, and of explaining the patterns of preserved and impaired capabilities of brain-damaged patients in terms of damage to one or more of the modules of a cognitive model. Since the process of identifying the functional lesion(s) in any individual case is *a posteriori,* the assumption of homogeneity of patients cannot be maintained; hence, the possibility of doing group studies involving patients is ruled out. Finally, classifying patients according to clinical syndromes contributes little to our quest for a better understanding of cognitive functioning.

2.0. Presence of Phonetic Information in Chinese Characters

The Chinese writing system, unlike alphabetic or syllabary writing systems, has no elements in a character that are consistently related to phonetic segments. Despite that approximately 90% of characters are phonetic compounds, many still consider the orthographic system in Chinese as non-phonetic (e.g., Coltheart (1984), Hoosain (1991)). Their view is based on the facts that (1) *not all* characters are phonetic compounds; there are orthographs that cannot be analyzed into a signific and a phonetic; (2) in cases where the signific in a phonetic compound is an existing graph and therefore has a pronunciation, a person who has never encountered the character before would not be able to decide which constituent is the phonetic; the situation is worse where there are radicals that can be a signific in some characters and a phonetic in others; (3) the phonetic is not a reliable indicator of the sound of characters, in the sense that the pronunciation of a phonetic compound is often different from that of its phonetic;

(4) even in cases where both the phonetic compound and its phonetic are phonologically similar in that they share the same segmental materials, they may still differ in tone; and finally (5) the frequency of occurrence of the phonetic compound as a whole can be higher than that of its phonetic as a character.

Although these are seemingly reasonable arguments, a closer examination reveals that they are in fact dubious. As mentioned earlier, phonetic compounds constitute the overwhelming majority of characters in the writing system. It is true that there are unanalyzable characters, but they are in the minority. It is also correct to point out that some radicals function as the signific in some characters and the phonetic in other characters; however, the position of occurrence of the phonetic component within a character is by no means equal in distribution. The strong tendency is that for characters with a left/right configuration, the phonetic occurs on the right, and for those characters with a top/bottom configuration, the phonetic appears at the bottom. Thus for a phonetic compound with two pronounceable parts, a reader can make a safe guess at which component is the phonetic, even without any prior exposure to the character. Even though many phonetic compounds are not identical in sound to their phonetic, statistics concerning the phonetic component does support its importance in providing phonological cue to characters containing them. According to Zhu (1987), there are 800 phonetics in Chinese. In modern day usage, about 26.5% of phonetic compounds in Mandarin have the same pronunciation as their phonetics. Using notions such as activation of entries from the information processing approach, phonetic compounds sharing the same phonetic component may be considered as belonging to the same activation region. Fang, Horng, and Tzeng (1986) surveyed 201 such regions, and found that in 18% of them, the pronunciation of the phonetic is identical to that of all the phonetic compounds in the same region. In addition, for regions that are associated with two or more phonological forms, 67% of the time the sound of the phonetic belongs to the dominant phonological group. Despite these statistics,one may still treat the phonetic component as an unreliable phonetic cue; however, considering the fact that in languages like English and French, the association between letters and phonemes is not always regular (that is, the mapping from letter to sound is often one to many), it is, therefore, not entirely reasonable to say that Chinese characters do not contain components that provide phonetic information about the characters.

The debate regarding the presence or absence of phonetic information in the Chinese orthographic system might continue if only the aforementioned observations were taken into account, but I believe the strongest argument for the claim that Chinese characters contain a great deal of phonetic information comes from the consistency effects obtained in the reading of Chinese characters and pseudo-characters (Fang, Horng, and Tzeng 1986; Lian 1985; Hue and Erickson 1988). These studies will be discussed in the next section.

2.1. Experimental Findings

The most important evidence demonstrating the role of the phonetic component in the reading of phonetic compounds is found in Fang, Horng, and Tzeng (1986), in which the naming latencies of phonetic compounds, real as well as pseudo-characters, with high and low consistency values were compared in a series of experiments. Consistency value indicates the extent to which characters sharing the same phonetic component (i.e., characters in the same activation region) are dominated by a phonological form. A high consistency value means that most of the characters in a region have the same pronunciation; whereas a low consistency value indicates that characters in a region are associated with two or more phonetic forms, without any one of them being the dominant form. For example, the activation region of the stem 巨 contains six lexical items (including the stem itself); they are all pronounced as *ju* (巨，拒，距，炬，苣，and 駏). Thus, the characters are said to have a high consistency value. In contrast, the activation region of the stem 由 consists of twelve characters; two of which are associated with the sound *di* (迪 and 笛), two with *xiou* (袖 and 岫), two with *zhou* (宙 and 軸), one with *chou* (抽), and five with *you* (由，油，鈾，柚，and 釉). In other words, this region is linked to five phonetic forms, and thus characters of this region are considered having a low consistency value.

In the first experiment, the naming latencies of three types of characters were compared --regular/consistent characters, which are characters whose pronunciation is identical to that of their phonetic and that the activation regions to which they belong are dominated by a single phonological form (i.e., characters with high consistency values); regular/inconsistent characters, which are characters that are phonetically identical to their phonetic component, but the regions they belong to are connected to more than one pronunication (i.e., orthographs with low consistency values); and exception characters, which are graphs that bear no phonological resemblance to their phonetic and have low consistency values (e.g. 肥，路，姓). It was found that "regular/consistent" characters were named significantly faster than the other two types of characters, and "regular/inconsistent" and "exception" characters did not differ from each other with respect to naming latency. To ensure that the longer latencies of "exception" characters were not the result of the conflict in pronunciation between the characters and their phonetic, a second experiment was conducted, in which simple characters (that are not analyzable into a signific and a phonetic) that may serve as the phonetic components in characters were selected. They were categorized into "high consistency value", "medium consistency value", and "low consistency value" groups. The consistency effects that were found in the first experiment were again observed. Low-consistency characters were named significantly slower than medium- and high-consistency characters; the latter two types did not differ significantly. To further illustrate the consistency effects, pseudo-characters were used in the third experiment. These stimuli were constructed by combining a high-frequency signific and a phonetic of either high-, medium-, or low-consistency value. The results paralleled those of the previous experiments. Pseudo-characters of high- and medium-consistency had significantly shorter naming latencies than those with low-consistency, and high- and medium-consistency pseudo-characters did not differ from each other. The findings from this set of experiments clearly demonstrate that the phonetic component plays an important role in reading Chinese phonetic compounds, despite the fact that these components do not necessarily have unique mapping with phonological forms.

That the phonetic is involved in oral naming of phonetic compounds finds support in other studies. Lian (1985) examined the effects of the level of consistency by comparing three types of characters-- "regular/consistent", regular/inconsistent", in the same sense as that in Fang, Horng, and Tzeng (1986), and "non-phonetic" characters, which are graphs that are not analyzable into a signific and a phonetic, and do *not* serve as the phonetic component in other characters. The results were consistent with those in Fang, Horng, and Tzeng's study, in that "regular/inconsistent" characters were named significantly more slowly than "regular/consistent"; "non-phonetic" orthographs had longer response time than "regular/consistent", but shorter than "regular/inconsistent", although the difference in reaction time between "non-phonetic" and "regular/inconsistent" characters,and that between "non-phonetic" and "regular/consistent" characters did not reach significance. In analyzing the naming errors in each type of stimuli, it was found that the "regular/inconsistent" group had the highest error rate. The majority of these errors were biased toward the pronunciation of the phonetic. On the other hand, the errors in the "non-phonetic" group tended to be visually related to the targets, and those in the "regular/consistent" group were mostly no responses. The general pattern of the results on naming latencies not only once again confirmed the role of the phonetic in reading, but also suggested that when a phonetic component is associated with more than one phonological form, the presence of the phonetic may in fact have an interfering effect on the naming of the character containing it, resulting in longer reaction time, and that a phonetic that is consistently linked to one sound may have a facilitative effect on reading. [2] Similar results were obtained in Hue and Erickson (1988), although the consistency effects were found only in characters of relatively low frequency of occurrence.

2.2. Neuropsychological Findings

This section presents a recent case study of a Mandarin Chinese brain-damaged patient (Tzeng, Hung, Lu, Chen, and Hu 1992), whose reading

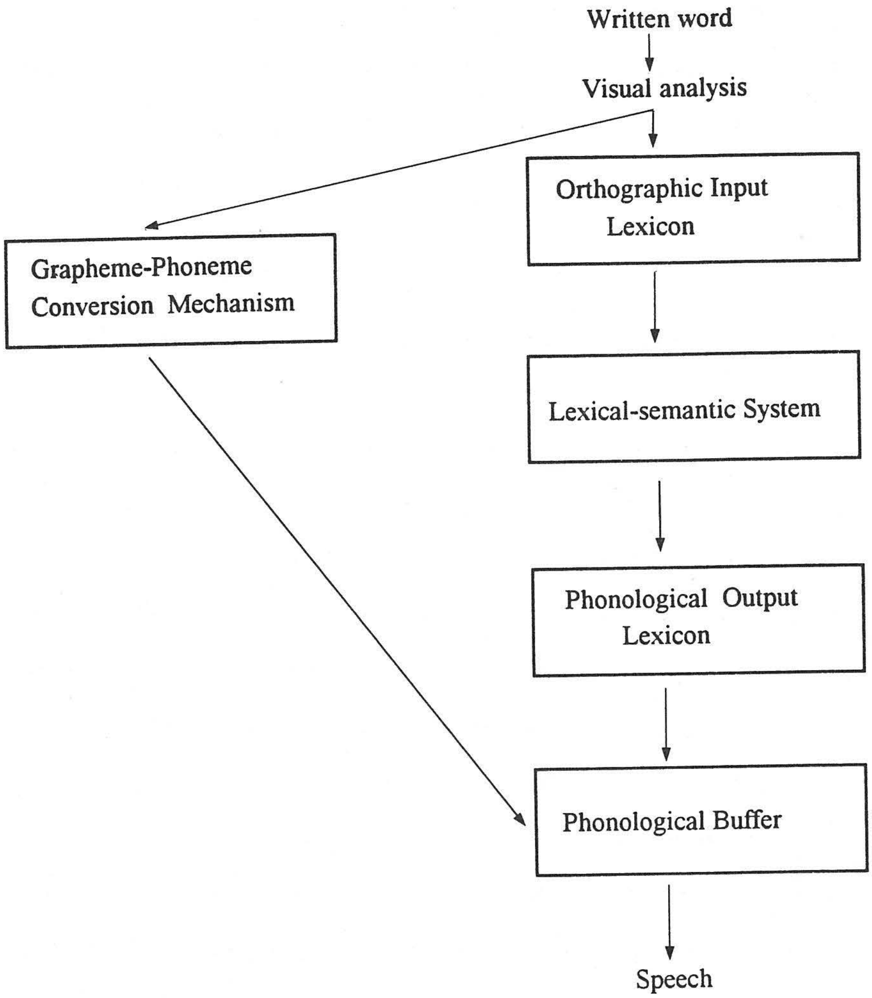
2 Although in Fang, Horng, and Tzeng (1986) and Lian (1985), the results were interpreted as evidence for an "activation--synthesis" model of reading proposed by Glushko (1979), they were not inconsistent with the dual--route model (Coltheart 1978; 1980).

performance calls into question the claim that the Chinese writing system is non-phonetic. However, before discussing this patient, I will first provide some theoretical background against which issues concerning the classification of the Chinese orthographic system are raised by the specific pattern of performance of the patient.

2.2.1. Lexical/Non-lexical Route of Reading

A very influential model of word processing based largely on data from alphabetic writing systems is the dual-route theory of reading (Coltheart 1978; 1980; Forster 1976; Morton and Patterson 1980). In this model, reading of words can be acheived via two functionally independent ways, i.e., lexical and non-lexical. To arrive at the pronunciation of a word by way of the lexical route, the target letter string first activates an entry in the orthographic input lexicon, which is a representation corresponding to the stimulus at the whole-word level, this entry will then activate a semantic representation of the presented item in the lexical-semantic system, which then activates a corresponding entry in the phonological output lexicon making available to the reader the spoken form of the word in question, as illustrated in Figure II.

Figure II. A dual route model of reading



The non-lexical route of reading derives the pronunciation of a letter string by first parsing the string into letters or letter groups, which then map onto phonemes via a set of grapheme-phoneme conversion rules. Since this mechanism operates at a level smaller than a word, it may be used to provide pronunciations for both words and invented non-words, such as *goom, wull.* The reading of the latter word type is assumed to be performed via the non-lexical route, because no corresponding representations of non-words are stored in the orthographic input lexicon. The success of the grapheme-phoneme conversion mechanism in arriving at the correct pronunciation of real words is dependent on the degree of regularity of the letter-sound correspondence in a particular lexical item. Presented with irregular words, such as *yacht, have,* or *listen,* whose letter-sound relation violates the usual grapheme-phoneme correspondence, the system will produce "regularization" errors, in this case, generating /yakt/, /hev/, /lɪstən/, respectively.

The postulation of the non-lexical route of reading is supported by the observation that normal readers can read not only existing words but also non-words or highly unfamiliar words which are unlikely to have representations in the orthographic input lexicon. [3] Another source of evidence comes from several types of reading disorders found in brain-injured patients. These disorders are believed to be the results of selective impairment to either the lexical or the non-lexical route. Patients who are classified as presenting with "phonological dyslexia" can read real words with ease but perform poorly on reading non-words (Beauvois and Dérousné 1979, Shallice and Warrington 1980). This pattern of impaired and preserved performance of reading is generally explained in terms of a damage to the non-lexical route. Patients who are categorized as "deep dyslexics" are also believed to have suffered from a damage to the non-lexical route. Their reading performance is characterized by their frequent production of semantic errors; for example, reading *act* as "play", *close* as "short" (Coltheart 1980; Coltheart, Patterson, and Marshall 1980; Marshall and Newcombe 1973; 1980). Other

3 For an alternative view regarding the reading of non-words, please refer to Glushko (1979) and Campbell (1983).

characteristics of this syndrome include very poor reading of non-words, a grammatical class effect in that content words are read better than function words (Saffran and Marin 1977), the occurrence of morphological errors (Patterson 1980; Shallice and Warrington 1975), visual errors (Coltheart, Patterson, and Marshall 1980), and an effect of imageability, such that nouns with high imageabilty are more likely to be read correctly than those with low imageabilty (Marshall and Newcombe 1973; Shallice and Warrington 1975). It has been argued that an impairment to the non-lexical route is a necessary condition for the occurrence of semantic errors and the poor reading of non-words in deep dyslexic patients (Morton and Patterson 1980; Newcombe and Marshall 1980; Shallice and Warrington 1980). In contrast with phonological dyslexics and deep dyslexics, patients who are classified as "surface dyslexics" are considered to have an impaired lexical route (Marshall and Newcombe 1973). Generally speaking, these patients are able to read non-words and real words with regular spelling-to-sound correspondence, and fail to read correctly irregular or exception words. Their reading errors consist mainly of regularization errors, as a result of applying grapheme-to-phoneme conversion rules to exception words (Shallice, Warrington, and McCarthy 1983). The preserved ability to name non-words and regular words, and the tendency to regularize exception words are accounted for in terms of an intact non-lexical route and an impairment along the lexical route, be it to the orthographic input lexicon, the lexical-semantic system, or the phonological output lexicon.

In short, the double dissociation between the ability of reading real words and that of reading non-words demonstrates the possibility of selective damage to one of the two routes of reading with preservation of the other, providing support to the dual-route model of word processing.

2.2.2. Data from a Chinese Brain-damaged Patient

The foregoing discussion shows that reading in alphabetic languages may be acheived via a lexical or a non-lexical route. Returning to the case of Chinese, if Chinese characters were basically pictographs and did not support any mapping between orthographic forms (below the level of word) and sounds, then it would be unlikely that a non-lexical route of reading Chinese graphs exist, and character naming can only be done through the lexical route. This means that pseudo-character naming would not be possible (which has already been shown in section 2.2.1. to be incorrect with normal subjects), and that the pronunciation of characters can only be arrived at post-lexically, i.e., an entry in the phonological output lexicon corresponding to the target character is activated by input from the lexical-semantic system. Put in the context of acquired reading disorders, Chinese patients with acquired aphasia would be predicted to be unable to read pseudo-characters, and if they could name a character, they must also have access to the meaning of the character. However, a recent report of a Chinese aphasic reader, H. C., showed that the predictions were not borne out (Tzeng, Hung, Lu, Chen, and Hu 1992).

H.C. is a 68-year-old male who was a teacher in a university in Taiwan. He suffered from a cortico-basal ganglionic degeneration. His performance in BDAE (Boston Diagnostic Aphasia Examination) showed good repetition and fluent spontaneous speech; however, his poor performance on object naming, reading and auditory comprehension indicated a general deterioration in comprehension. Despite his impaired performance on all comprehension tasks, he was able to read rapidly and effortlessly characters varying in degree of concreteness, frequency of usage, and the form in which they were written (simplified as well as traditional form). Although his naming of these orthographs was spontaneous, he did not seem to know their meaning, as evidenced by his poor performance in a character-picture matching task, and a categorization task involving sorting characters which were names of animals, fruits, or articles of everyday use in terms of semantic category. In other words, he was able to arrive at the correct pronunciation of characters the meaning of which he had no knowledge. [4] Furthermore, when he was asked to name pseudo-characters, he would produce a pronunciation on the basis of the phonetic component in the made-up character. Interestingly, he didn't seem to be able to distinguish between pseudo-characters and real characters; on some occasions, he even fabricated meanings for the made-up characters, based on the associated meaning of their signific components. In sum, the facts that H. C. could read characters correctly with little knowledge of their meaning, and that he could pronounce pseudo-characters by extracting information from their phonetic component, argue against the claim that Chinese characters do not contain phonological information.

Conclusion

The Chinese writing system has been considered a deep orthography, since the character-sound mapping appears to be very opaque; as a consequence, the system is believed by many to be non-phonetic in nature. However, this view is seriously challenged by the reading performance of normal subjects and the pattern of intact and impaired language performance

4 This may be taken to imply the existence of a direct route from the orthographic input lexicon to the phonological output lexicon in Chinese. Such a route has been proposed for English (Funnell 1983).

of a brain-damaged individual. The studies discussed in section 2.1. clearly demonstrated a consistency effect in reading phonetic compounds and pseudo-characters composed of a signific and a phonetic. Characters, real and invented, with phonetic components that are high on consistency value were named faster than ones with low-consistency phonetics. The expectations, stemming from the non-phonetic view of Chinese orthographs, that character naming is achieved post-lexically, and that Chinese cannot read pseudo-characters, were disconfirmed by the report of one brain-injured patient, who could read both real and pseudo-characters effortlessly with little comprehension of these graphs. These results not only point to the importance of the phonetic component in reading, but also provide valuable insights into the nature of input orthographic representations of Chinese characters. That is, if phonetic compounds are automatically parsed into components in reading, with one of them corresponding to the phonetic, then input orthographic representations of these characters must have internal structure. [5]

3.0. Internal Structure of Chinese Characters in Writing

On the basis of findings from experimental studies with normal subjects and from neuropathology, the claim that the Chinese writing system is non-phonetic was shown to be incorrect. In fact, these findings argue strongly that Chinese characters contain considerable phonological information. This section will further demonstrate that Chinese orthographs are not pictographs; on the contrary, they have internal structure. This is inferred from the types

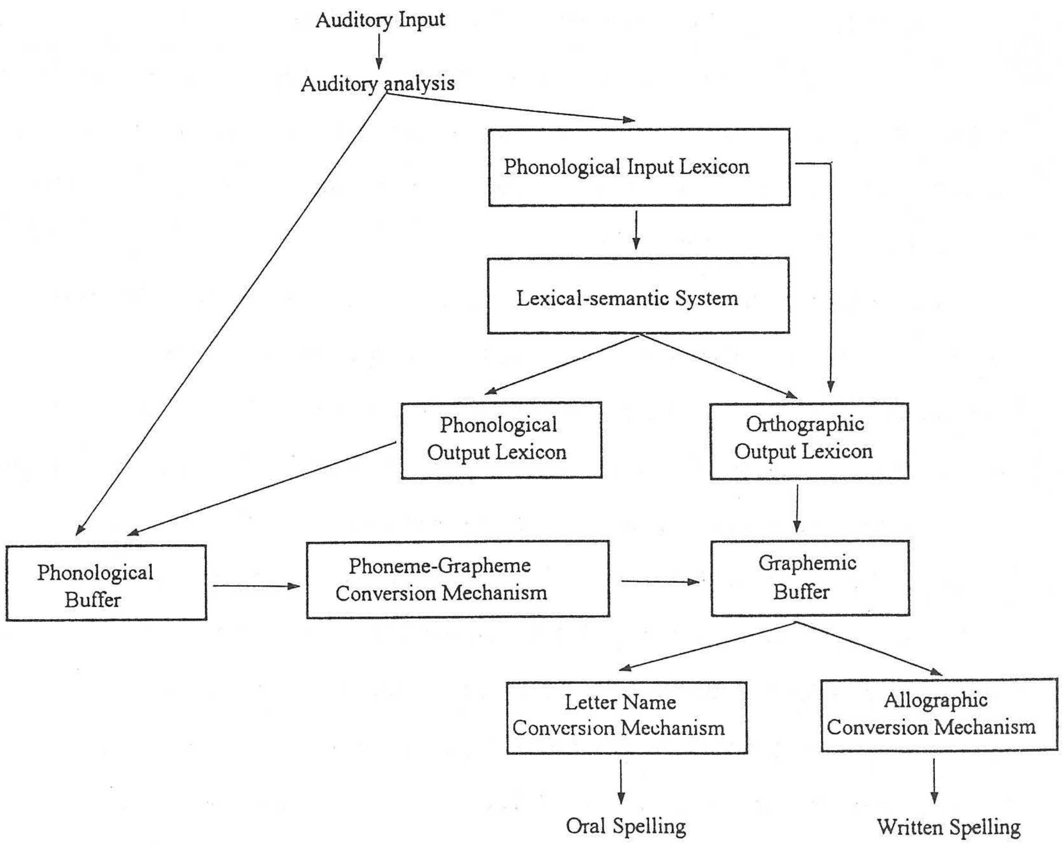
5 Given the available data, further specification for the nature of the internal components of input orthographic representations is not possible at this point.

of writing errors produced by Chinese dysgraphic patients, and the orthographic components involved in these errors. The reasoning is that if certain components can be manipulated independently, they must correspond to units in the orthographic representations of characters. Before presenting the proposal of the orthographic structure of Chinese characters in writing, I will first discuss an analysis of the graphemic representations of English and Italian words from the work of Caramazza and Miceli, which inspires the current proposal.

3.1. Structure of Output Graphemic Representations in English and Italian

In Caramazza, Miceli, Villa, and Romani (1987), Caramazza and Miceli (1990), and McCloskey, Goodman-Schulman, and Aliminosa (1990), analyses of the structure of output graphemic representations of English and Italian words were put forth. Output graphemic representations are forms based on which oral or written output of words are produced, upon appropriate conversion procedures, i.e., letter-name conversion for oral spelling and allographic conversion for written spelling. Figure III shows a schematic representation of the functional architecture of the spelling process.

Figure III. Functional organization of the spelling system



The term "grapheme" refers to abstract letters that are associated with single phonemes. For example, the word *tip* is said to contain the graphemes <t>, <i> , and <p>, which respectively map onto the phonemes /t/, /ɪ/, and /p/. The graphemic buffer in Figure III is a working memory system which temporarily stores a graphemic representation for letter-name conversion or allographic conversion, depending on the output modality. The diagram also indicates that the graphemic representations that are temporarily maintained in the graphemic buffer can be information about lexical items from the orthographic output lexicon, or the product of phoneme-grapheme conversion mechanism in the case of non-word spelling.

Borrowing heavily from autosegmental phonology, Caramazza and Miceli (1990) suggested that graphemic representations also have a multi-dimensional structure, with each dimension representing independent information. The graphemic structures of three Italian words, *tavolo* (table), *scambi* (exchanges), and *stella* (star) (from Caramazza and Miceli 1990, p. 282) are given in Figure IV.

The representations shown in Figure IV contain several dimensions (or tiers) -- one specifying the identities of graphemes comprising a word (grapheme tier), one representing the CV status of individual graphemes (CV tier), one indicating the quantity of the graphemes in question (quantity tier), and one representing the graphosyllabic boundaries within a word (graphosyllabic tier). Units in different tiers are linked by association lines, such that each unit of computation consists of information about grapheme identity-quantity-CV status-graphosyllabic structure. The postulation of each dimension is motivated by different types of writing errors from dysgraphic patients and the specific nature of the errors as a function of lexical characteristics, such as the presence or absence of geminates and the graphosyllabic structure of words. Evidence providing support for each tier as discussed in Caramazza and Miceli (1990) based on an extensive and detailed analysis of writing errors from an Italian dysgraphic patient, L. B., is briefly presented below. It is assumed that forms of errors may reflect the selective impairment(s) to one or more components of a computational system. An impairment may sometimes be so specific that only a certain informational dimension is affected, resulting in a highly constrained pattern of performance. To the extent that a proposed structural representation successfully accounts for the characteristics of a set of data, the data are then considered

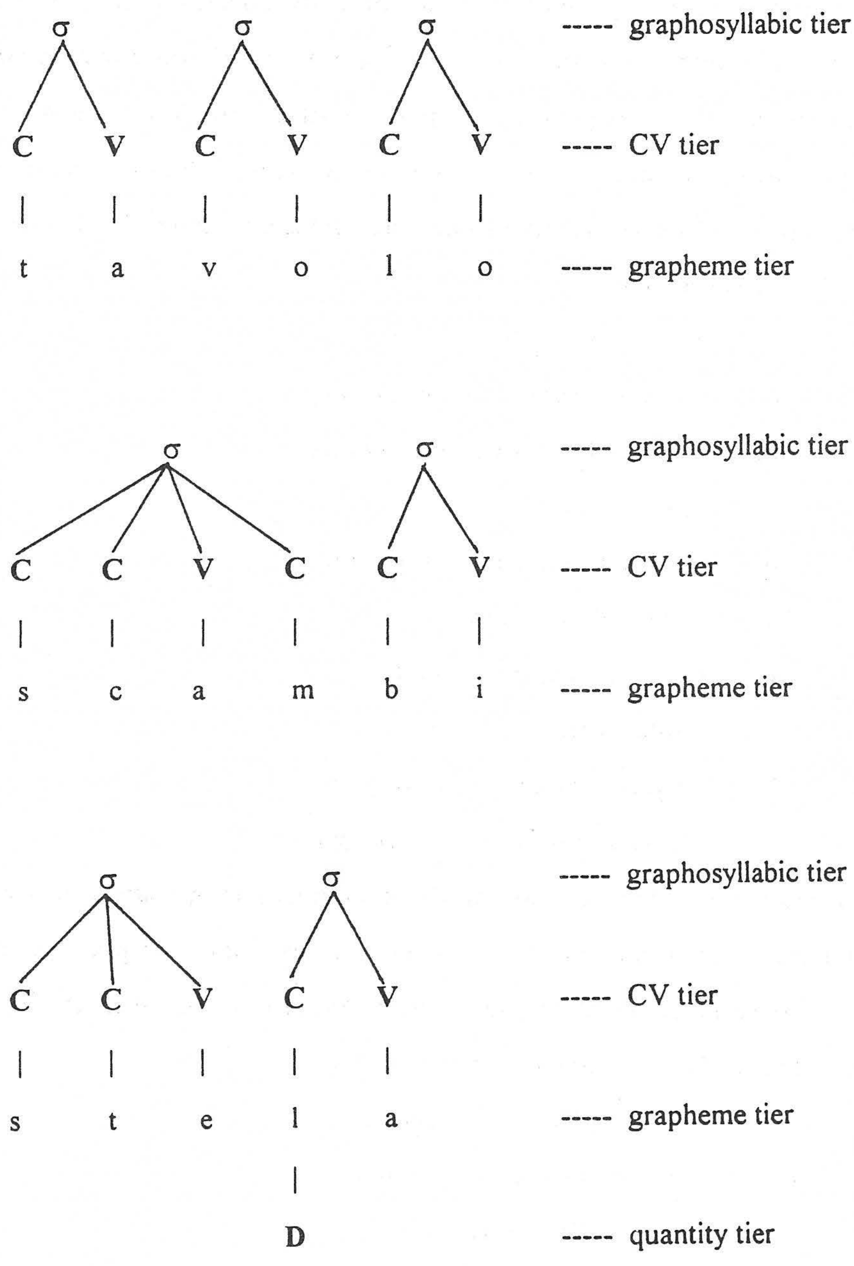
Figure IV. Graphemic representations of three Italian words:

tavolo, *scambi*, and *stella*

i.

ii.

iii.



supportive of the proposed analysis.

The specification of graphemes and their relative ordering are the most basic elements in a graphemic representation. If the identity of constituent graphemes in a representation is not fully specified, one of two things may happen -- either the underspecified grapheme(s) are substituted for by other graphemes, or they are simply omitted in the output, resulting in substitution errors in the former case and deletion errors in the latter. Examples of letter substitutions and deletions from L. B. are given in (1) and (2), respectively.

(1) Letter substitution errors

i. enorme (huge) ---> emorme

ii. metodo (method) ---> metoto

iii. favola (tale) ---> favala

iv. nucleo (nucleus) ---> nuclio

(2) Letter deletion errors

i. scatto (rush) ---> _catto

ii. svolta (turns) ----> svol_a

iii. volgare (vulgar) ---> vol__re

iv. albergo (hotel) ---> a_be_go

In cases where the content of the graphemes in question is preserved but the positions of some of the graphemes are not adequately specified, letter exchanges or letter shifts may occur. Examples are given in (3).

(3) a Letter exchange errors

i. sirena (siren) ---> serina

ii. secolo (century) ---> seloco

b Letter shift errors

i. giusto (right) ---> gusito

ii. potenti (powerful) ---> ponteti

Errors involving the addition of a letter to the target letter string were also found, although the source of these errors is not yet well understood. Some examples are given in (4).

(4) Letter addition errors

i. gancio (hook) ---> giancio

ii. cenere (ashes) ---> cencere

iii. binario (track) ---> biniario

iv. divano (sofa) ---> divarno

In addition to information about the position and the identity of individual graphemes, the representations in Figure IV also contain specifications for the CV status, the quantity of component graphemes, and the graphosyllabic boundaries of the letter strings. The postulation for the CV dimension was motivated by the observation obtained from L.B. that 99.3% of the letter substitution errors in the corpus respected the consonant/vowel distinction, and that the great majority of letter exchange errors involved consonants with consonants or vowels with vowels, as illustrated by examples in (1) and (3)a, respectively. A consonant/vowel distinction in substitutions and transpositions argues that CV information must be present in graphemic representations; furthermore, the information must be represented independently of that about the content of graphemes. Under such an assumption, a proper selection for a substitute in terms of CV status is possible despite loss of grapheme identities. In the case of letter exchanges, only letters having the same CV status can undergo transposition without resulting in a mismatch in mapping between the CV tier and grapheme tier. That the specification for the CV status of constituent graphemes is part of the structure of graphemic representations also found support in McCloskey, Goodman-Schulman, and

Aliminosa (1990).

Referring to the graphemic representation of *stella* in Figure IV, the geminate is represented by a single grapheme <l> (as opposed to two graphemes), marked by a feature D indicating that the grapheme in question is associated with two letters of the same content in oral and written spelling. This structural representation imposes severe constraints on the possible forms of errors that may occur in words containing geminates. First, if the to-be-doubled grapheme is substituted, the substitution should be realized in spelling as a substitution of the geminate cluster and not just one member of the geminate; second, if the quantity feature is somehow dissociated from the to-be-doubled grapheme and becomes reattached to another grapheme (i.e. a shift of geminate feature), the resultant form will contain a geminate involving a letter other than the target one; third, in cases where a graphemic representation is deformed in such a way that the positions of the to-be-doubled grapheme and another grapheme of the same CV status are not sufficiently specified, giving rise to an exchange error, the result will be a transposition between the geminate and the single letter. Examples illustrating different error types involving words with a geminate are given in the left-hand column in (5); forms of errors that are not predicted to occur and indeed were not found in L.B.'s extensive data corpus are listed in the right-hand column.

(5) Examples of errors on geminate clusters

i. Substitution errors

marrone (brown) ---> mazzone *marrone ---> marzone

ii. Shift errors

avviso (notice) ---> avisso *avviso ---> avivso

iii. Exchange errors

sorella (sister) ---> sollera *sorella ---> solelra

Two other predictions can be made based on the current proposal of graphemic representations with a quantity tier. On the assumption that addition of association lines linking the feature D with graphemes other than the to-be-doubled grapheme is possible, geminate duplication is predicted to occur in lexical items already containing a geminate, and rarely in ones without a geminate. This prediction was borne out, as observed in L.B.'s error corpus-- only 10 errors involved the production of a geminate in a corpus consisting of spelling 4400 words that do not have geminates. In addition, in the proposal by Caramazza and Miceli, a geminate cluster is represented by a grapheme marked by a feature D, rather than by two graphemes of the same content. Under the assumption that it is the number of grapheme identities and not the number of letters that can affect a subject's performance in spelling, given a subject who spells shorter words better than longer words, it is expected that words having a geminate should be spelled better than words of the same length in letters but without a geminate. This was found in the data from L. B. who showed a word length effect. Six-letter words containing a geminate cluster were spelled correctly 74% of the time, as opposed to 56% correct for 6-letter words without a geminate.

The claim that graphemic representations contain graphosyllabic information is supported by several pieces of observation from L.B.'s data. It was found that deletions of consonant or vowel rarely occured with consonants flanked by vowels, or vowels flanked by consonants; vowel or consonant deletions always took place in vowel or consonant clusters. This can be explained if one assumes that the simplest graphosyllable has the structure of a CV sequence, and appeals to the minimum complexity principle, which says that in response to a damaged representation, produce the least complex graphosyllabic sequence consistent with the available information. Hence, deletions of consonants surrounded by vowels would result in a more complex graphosyllable; on the other hand, deletions of consonants in consonant clusters would produce a less complex graphosyllable. The same situation is true of vowel deletions. It was also found that the distribution of errors involving tautosyllabic CC clusters (e.g., the CC cluster <pr> in a/pri/re (open)) differed qualitatively from that involving heterosyllabic CC clusters (e.g., the CC cluster <lp> in al/pi/no (alpine)). More specifically, the error rates of constituent consonants in tautosyllabic CC clusters were found to be comparable; whereas more errors were made on the first consonant in a heterosyllabic CC clusters (i.e., the coda of a graphosyllable) than on the second consonant of such clusters (i.e., the onset of a graphosyllable). This difference in distribution of error rate over members of tautosyllabic versus heterosyllabic CC clusters can be explained on the analysis that graphosyllabic boundaries are represented independently of other information in graphemic representations. The third observation relevant to the present proposal regarding graphosyllabic structure is that there were thirteen adjacent letter exchanges involving a vowel and a consonant in the corpus; all of them took place within a graphosyllable (e.g., scelta (choice) ---> scleta; figlia (daughter) ---> figila). This would make sense if it is assumed that the mechanism (in this case, the allographic conversion mechanism) converting graphemic representations temporarily stored in the graphemic buffer into letter strings for written spelling operates with a window of the size of approximately one graphosyllable.

Summary

On the basis of the fact that L. B.'s writing errors consisted of letter substitutions, deletions, insertions, shifts and transpositions, his functional lesion was hypothesized to be at a post-lexical level. [6] In addition, a highly elaborate multi-dimensional structure of graphemic representations containing information of various kinds (i.e., grapheme identity, position of occurrence, CV status, quantity, and graphosyllabic structure) was proposed. This was made possible through a detailed analysis of the specific nature of L.B.'s errors. The previous discussion, therefore, illustrates how data from brain-damaged patients can inform and constrain models of the cognitive system.

3.2. Structure of Output Orthographic Representations of Chinese Phonetic Compounds

In this section, a proposal about the structure of output orthographic representations of Chinese phonetic compounds is presented. As a starting point, it is assumed that the structure of the lexical system in Chinese has the form as depicted in Figure I. The proposed representations of phonetic compounds contain the identity and the position of the signific and the phonetic. Evidence supporting this proposal comes from writing errors of

6 In fact, the locus of damage in L. B.'s word processing system was further specified to be in the graphemic buffer, which receives input of graphemic representations from the orthographic output lexicon through lexical retrieval, or from the phoneme-grapheme conversion mechanism through the application of phoneme-grapheme association rules. The conclusion concerning the identity of the functional lesion in question was made based on the observations that the spelling of words and non-words were comparably affected, that both oral and written spelling were impaired, and that an effect of word length was found.

three Cantonese Chinese dysgraphic patients in Hong Kong. Since these errors are preliminary data of a project investigating the processing system of single Chinese words, only a general level of description of the structure is possible at this point. Despite the degree of generality in description, the claim that Chinese characters have internal structure constitutes an argument against the view that Chinese orthographs are pictographic in nature.

The three dysgraphic patients whose writing errors are discussed here include two male and one female native speakers of Cantonese Chinese. The age ranged from 38 to 49. They all had at least secondary school level of education. Their former professions suggest that they were premorbidly good writers.

Table I. Information about three Cantonese Chinese dysgraphic patients in a preliminary study investigating the cognitive processes of single Chinese words

	D. G.	Y. K.	C. L.
Age	38	49	41
Gender	F	M	M
Handedness	Right	Right	Right
Former profession	policewoman	journalist	head of a department of a major airline
Etiology	subarachnoid haemorrhage caused by aneurysm at the left carotid artery	subarachnoid haemorrhage with hydrocephalus caused by aneurysm at the right antero-communicatory artery	left lateral cerebral haemorrhage, craniotomy done

On the basis that these patients could repeat single characters and phrases perfectly, and they did not show signs of comprehension impairment, it is assumed that their errors arose from a post-lexical level. In other words, their errors were not the result of incorrect activation of entries in the phonological input lexicon due to misperception of the stimuli, or the result of selection errors in the orthographic output lexicon due to damage to the lexical-semantic system.

Recall that the postulation of a grapheme tier consisting of an ordered letter string in the orthographic representations of English and Italian words was motivated by the existence of letter substitution, letter deletion, letter insertion, and letter transposition errors. Similarly for the case of Chinese phonetic compounds, if these types of errors often refer to certain components in a character, then it is reasonable to suggest that such components correspond to constituents in orthographic representations, comparable to graphemes in graphemic representations. In the rest of this section, I will present writing errors of these types from the writing to dictation task and the written picture naming task performed by the patients in our preliminary study. Since the signific and the phonetic were found to be involved in these error types, it is argued that the orthographic representations of phonetic compounds must at least contain information concerning the content and the position of occurrence of the signific and the phonetic component.

Substitution errors

If the orthographic representation of a phonetic compound is deformed in such a way that the identity of either the signific or the phonetic is not sufficiently specified, substitution errors involving these elements may occur. Examples (6) and (7) illustrate signific component substitutions and phonetic component substitutions, respectively. In (6)b, the rightmost column provides the pronunciations of the target character (the one on the left) and the output (the one on the right); whereas in (7)a and b, the rightmost column gives the phonetic forms of the phonetic components (in the target and its response) when they occur as characters. The numeric values next to the segmental materials represent the tones.

(6) Signific component substitutions

a. Non-existent character response

Target	Response
貓 (cat)	⿰虫苗
踢 (to kick)	揚
眺 (to look afar)	⿰耳兆
跌 (to fall)	蛈
鈕 (button)	杻
螺 (conch)	⿰目累
癌 (cancer)	⿸广嵒
筆 (pen)	⿱艹聿
風箏 (kite)	風⿱艹爭
麓 (foot of a hill)	鹿

b. Real character response

Target	Response	Contrast in pronunciation
爬 (to crawl)	把 (to take)	[pʰa] 11 vs. [pa] 35
瘀 (bruise)	淤 (to clot)	[y] 35 vs. [y] 55
柏 (pine)	泊 (to dock)	[pʰak] 33 & [pʰak] 33
蛙 (frog)	娃 (doll)	[wa] 55 & [wa] 55
炒 (to stir-fry)	抄 (to copy by hand)	[tʃʰau] 55 & [tʃʰau] 55
蜢 (grasshopper)	猛 (fierce)	[maŋ] 35 vs. [maŋ] 13
籃 (basket)	藍 (blue)	[lam] 35 vs. [lam] 11
駕 (to drive)	架 (shelf)	[ka] 33 vs. [ka] 35
籃 (basket)	艦 (warship)	[lam] 35 vs. [lam] 22
茄 (2nd character in cigar)	架 (shelf)	[ka] 55 vs. [ka] 35
彰 (to display)	障 (obstacle)	[tʃoeŋ] 55 vs. [tʃoeŋ] 33

(7) Phonetic component substitutions

a. Non-existent character (or word) response

Target	Response	Phonetic component contrast
煮 (to cook, [tʃy] 35)	暑	[tʃɛ] 35 vs. [sy] 35
盅 (container, [tʃoŋ] 55)	盅	[tʃoŋ] 55 vs. [tʃʰyn] 33
塔 (tower, [tʰap] 33)	塔	non-existent as character vs. [tap] 33
爺 (grandfather, [jɛ] 11)	爺	[jɛ] 11 vs. [tʃʰøy] 35
抓 (to scratch, [tʃau] 35)	抓	[tʃau] 35 vs. [kwa] 55
財 (fortune, [tʃʰɔi] 11)	財	[tʃʰɔi] 11 vs. [tʃʰyn] 33
霉 (moldy, [mui] 11)	霉	[mui] 13 vs. [mo] 13
菇 (fungus, [gu] 55)	菰	[gu] 55 vs. [wu] 11
喇叭 (trumpet, [la] 33 [pa] 55)	喇 吧	[pat] 33 vs. [pa] 55

b. Real character response

Target	Response	Phonetic component contrast
盅 (container, [tʃoŋ] 55)	盃 (cup/medal, [pui] 55)	[tʃoŋ] 55 vs. [pɐt] 55
盅 (container, [tʃoŋ] 55)	盔 (helmet, [kwʰɐi] 55)	[tʃoŋ] 55 vs. [fui] 55
胸 (chest, [hoŋ] 55)	腔 (cavity, [hɔŋ] 55)	[hoŋ] 55 & [hoŋ] 55
蟬 (cicada, [sim] 11)	蟾 (1st character in toad, [sim] 11)	[tan] 55 vs. [sim] 55

The errors listed in (6)a are substitutions of the signific component; the resultant forms are non-existent characters. Three observations are made: (i) the element substituting for the target is also a signific; (ii) the substitute occupies the same position as the target signific, except for the last error 麓 (foot of a hill); and (iii) the substitute appears in a position where it normally would when it functions as a signific in other characters. One may notice that the majority of significs in question occur in the left half of the stimuli,and two appear on the top. This simply reflects the strong tendency that the signific occupies the left side in characters with left/right configuration, and on the top in graphs with top/bottom arrangement. The three characteristics just mentioned argue for two things -- the signific component must correspond to a constituent at some level of processing in the orthographic representations for it to be replaced as a unit, and the signific (both the target and the substitute) must be spatially coded so as to explain the observations that the substitute appears in the same position as the target signific, and that it occupies its usual position when it is a signific in other characters. The responses in (6)b are real characters. In these cases, it is always possible to interpret these errors as lexical substitutions. One thing worth noting is that except for the last three errors 籃 (basket), 茄 (as in 雪茄 (cigar)), and 彰 (to display), the target and the response in each case have their signific component occurring in the same position. [7]

The errors in (7)a illustrate that when the information about the identity of the phonetic component in an orthographic representation of a phonetic compound is underspecified, phonetic substitutions may be resulted.

7 Notice that the phonological form of the target and that of its response are highly similar, if not identical, as shown in the rightmost column in (6)b. Except for the stimulus 爬 (to crawl), the forms in question differ only by tone. Thus, one may argue that the reponses were phonologically plausible errors resulted from insufficient activation of the appropriate entries in the orthographic output lexicon received from the lexical-semantic system, and activation of representations in the orthographic output lexicon that are phonologically similar to the targets. Although the available data do not allow us to rule out such a possibility, it is unlikely to be the case on the basis that these patients did not show any perceptual problems or comprehension difficulties. The same reasoning applies to other errors with real character responses to be discussed in the rest of the paper.

Except for 煮 (to cook) and 菇 (fungus), the elements substituting for the target phonetics function as phonetics in other characters. The errors in (7)a show that the selection of a substitute phonetic is not necessarily related to the phonological similarity between the target phonetic and the substitute phonetic, as indicated by the pairwise contrasts in the right-hand column. One may also notice that the stimuli and the responses in (7)a are visually similar, except for the target 菇. It is thus arguable that the errors are in fact the results of substitution, deletion, or addition of some part not corresponding to the phonetic component. More specifically, it is plausible that 煮 and 蛊 may, in fact, involve the addition of 日 and 口, respectively; 塔 and 爺 involve the substitution of ⺮ for 艹, and 又 for 阝, respectively; 抓 and 財 are the results of stroke substitutions, such that 爪 becomes 瓜, and 才 becomes 寸, respectively; finally, 霉 is the product of the deletion of 𠂉. Although such possibilities cannot be ruled out, the fact that in all these errors, the resultant forms contain a signific and a remainder which is an existing character suggests that the responses were more likely to be the products of phonetic substitution. The responses in (7)b are real characters; analogous to the situation in (6)b, one cannot be certain whether these responses were lexical substitutions or the results of replacing the underspecified target phonetic with a substitute whose combination with the remainder of the character in question inadvertantly formed an existing character. Finally, in all the errors listed in (7), the substituting elements occupy the same position in a character as their target counterpart.

Deletion errors

Underspecification of the identity of the signific or the phonetic component may also lead to deletion errors. Examples of signific deletions and phonetic deletions are given in (8) and (9), respectively. The deletion of the signific component in phonetic compounds often results in real characters, for the fact that the great majority of phonetic components have independent existence. [8] Because of this, one cannot rule out the possibility of lexical substitutions. Less ambiguous signific deletion errors would involve phonetic compounds whose pronunciation differs from that of their phonetic (e.g., 跌 (to fall, [tit] 33) ---> 失 (to lose, [sɐt] 55)), or phonetic compounds whose phonetic does not exist alone (e.g., 僕 (servant, [pok] 22) ---> 菐). The errors in (9) illustrate the omission of the phonetic component. Although the first four errors contain significs that have independent existence (i.e., 虫 (worm) in 蜢 , 足 (foot) in 趾 , 貝 (shell or precious objects)in 賬 , and 火 (fire) in 煙), the manner in which they were written clearly show that the subjects intended them to be the significs only. In particular, the parts were written relatively small in comparison with whole characters produced by the subjects, and these components occupied the same position as they would appear in the targets.

(8) Signific component deletions

Target	Response
茄 (squash)	加 (to add)
菠 (spinach)	波 (wave)
菇 (fungus)	姑 (aunt)
芽 (bean sprout)	牙 (tooth)
蝴 (1st character in butterfly)	胡 (a last name)

8 Lian (1985) surveyed a total of 280 phonetic components and found that only 6 phonetics do not exist as characters.

(9) Phonetic component deletions

Target	Response
蜢 (grasshopper)	虫
趾 (toe)	⻊
賬 (itemized bill)	貝
煙 (cigarette)	火
浮 (to float)	氵
猴 (monkey)	犭

Addition errors

Several errors involving the insertion of a signific were found in the corpus; they are listed in (10). A signific 木 related to the meaning of wood/plant was added to the character 摩 (to rub), resulting in a non-existent graph. The second error was the product of a signific substitution (蜢---> 猛) and an addition of the signific 艹. For the third error, the character of interest was the second syllable of an adverbial meaning "just". Although the resultant graph is also a character, 材 (material), the combination of this graph with the first character of the target phrase is not a meaningful whole; in other words, it is a non-word. The responses in the last two errors are existent orthographs; they are also phonetically similar to their target. It is worth noting that no addition errors involving the phonetic were found. This may be due to the fact that the insertion of a phonetic component into a graph that already has a phonetic would result in a grossly ill-formed orthograph. The significance of this, if the reasoning is correct, is that not only is the production of errors highly constrained (meaning that errors are not generated randomly), the output may be subject to the filtering of orthographic well-formedness conditions, in this case, the condition that a character cannot have more than one phonetic component. One final note, although we do seem to have a reasonable account for the source of signific/ phonetic substitutions and deletions, we do not yet understand how addition errors arise. Nevertheless, the important point concerning the current proposal of the structure of orthographic representations is that the inserted elements must correspond to some units or constituents in the representations, such that when orthographic representations are deformed as a result of damage to the system, these units or constituents may be inserted as wholes.

(10) Signific component additions

Target	Response
摩 (to rub)	
蜢 (grasshopper)	
剛才 (just)	剛材
奉 (to offer, [foŋ] 22)	俸 (salary, reward, [foŋ] 35)
羊 (sheep, [jœŋ] 11)	洋 (ocean, [jœŋ] 11)

Transposition errors

If information about the position of the constituents (i.e. the signific and the phonetic) in a representation is not adequately preserved, transposition errors may be resulted, analogous to the letter exchange errors found in L.B.'s data discussed in section 3.1. (Caramazza et al. 1987; Caramazza and Miceli 1990). However, there were no transposition errors in the preliminary data of this study. [9] In fact, examining the data reported in previous studies

9 There were several transposition errors involving the exchanges of orthographic parts that are larger than a stroke but are not considered the signific or the

(Huang 1984; Li, Hu, Zhu, and Sun 1984; Lyman, Kwan, and Chao 1938; Naeser and Chan 1980; Tseng, Hung, Chen, Wu, and Hsi 1986), only one transposition error was listed, as shown in (11). The rarity of this type of errors may be due to the fact that a transposition error occurs when only the spatial information of the elements involved is lost, but their content is still retained. In the case of a phonetic compound, so long as the spatial code of either the phonetic or the signific is preserved, the position of occurrence of the other element is greatly constrained. This means that correct production of a phonetic compound is possible despite incomplete spatial information of the constituents.

(11) A transposition error from Lyman, Kwan, and Chao (1938)

期 (a period of time) ----> 朞

Summary

This section discussed the preliminary data of a study investigating the writing processes of single Chinese words. The corpus contained writing errors from three Cantonese dysgraphic patients. The different types of errors support the proposal that the output orthographic representations of phonetic compounds contain specifications for the identity and the position of occurrence of the signific and the phonetic component. This claim is based on the observations that substitution, deletion, and addition errors could refer to units corresponding to the signific or the phonetic, and that the substitute appeared in the same position as the substituted element.

---

phonetic of the characters. This does not necessarily undermine the proposal here. If anything, it points to the need for a thorough analysis of the basic constituents in the Chinese writing system in order to arrive at a proposal of the structure of output orthographic representations comparable in degree of elaboration with the one put forth by Caramazza and Miceli (1990).

4.0. Conclusion

The consistency effect of the phonetic component obtained from normal subjects' reading of phonetic compounds and pseudo-characters composed of a phonetic and a signific, the case report of a Mandarin Chinese brain-damaged patient who demonstrated a preserved ability in reading pseudo-characters and real characters with little knowledge of their meaning, together with the finding from Chinese dysgraphic patients that writing errors of the types substitution, deletion, insertion, and transposition could involve the signific or the phonetic, constitute support for the claim that Chinese orthographs indeed have internal structure. The evidence also renders the non-phonetic and pictographic view of the Chinese writing system untenable. The use of neuropathological data from reading and writing of single characters to draw conclusions about the structure of orthographic representations exemplies the way in which data and theories are related in the cognitive neuropsychological approach.

Having demonstrated the importance of the phonetic component in reading, one naturally wonders whether the signific also has a role to play in the processing of characters, and if it does, what is the nature of that role. It is conceivable that the signific component is involved in the comprehension of characters. In terms of writing, future investigations into the basic orthographic constituents of the Chinese writing system will allow us to articulate in greater detail the structure of output orthographic representations. Data from brain-injured patients will unquestionably shed light on these issues.

(Accepted for publication 19 August 1993)

Bibliography

連韻文

1985 中文唸字歷程的探討：聲旁的語音觸發作用．國立臺灣大學心理學研究所碩士論文．

Allport, D.A. and Funnell, Elaine

1981 Components of the Mental Lexicon, *Philosophical Transactions of the Royal Society (London),* B295: 397-410.

Badecker, William and Caramazza, Alfonso

1985 On Considerations of Method and Theory Governing the Use of Clinical Categories in Neurolinguistics and Cognitive Neuropsychology: The Case against Agrammatism, *Cognition,* 20: 97-125.

Bates, Elizabeth, Appelbaum, Mark, and Allard, Lee

1991 Statistical Constraints on the Use of Single Case in Neuropsychological Research, *Brain and Language,* 440: 295-329.

Bates, Elizabeth, McDonald, Janet, MacWhinney, Brian, and Appelbaum, Mark

1991 A Maximum Likelihood Procedure for the Analysis of Group and Individual Data in Aphasia Research, *Brain and Language,* 40: 231-265.

Beauvois, M.F., and Dérousné, J.

1979 Phonological Alexia: Three Dissociations, *Journal of Neurology, Neurosurgery and Psychiatry,* 42: 1115-1124.

Campbell, Ruth

1983 Writing Nonwords to Dictation, *Brain and Language,* 19: 153-178.

1987 One or Two Lexicons for Reading and Writing Words; Can Misspelling Shed Any Light, *Cognitive Neuropsychology,* 4: 487-499.

Caramazza, Alfonso, Miceli, Gabriele, Villa, Giampiero, and Romani, Cristina

1987 The Role of the Graphemic Buffer in Spelling: Evidence from a Case of Acquired Dygraphia, *Cognition,* 26: 59-85.

Caramazza, Alfonso and Miceli, Gabriele

1990 The Structure of Graphemic Representations, *Cognition,* 37: 243-297.

Caramazza, Alfonso

1984 The Logic of Neuropsychological Research and the Problem of Patient Classification in Aphasia, *Brain and Language,* 21: 9-20.

1986 On Drawing Inferences about the Structure of Normal Cognitive Systems from the Analysis of Patterns of Impaired Performance: The Case for Single-Patient Studies, *Brain and Cognition,* 5: 41-66.

1988 Some Aspects of Language Processing Revealed through the Analysis of Acquired Aphasia: The Lexical System, *Annual Review of Neuroscience,* 11: 395-421.

1992 Is Cognitive Neuropsychology Possible? *Journal of Cognitive Neuroscience,* 4. 1: 80-95.

Coltheart, Max

1980 The Semantic Error: Types and Theories, *Deep dyslexias,* London: Routledge and Kegan Paul.

1984 Writing Systems and Reading Disorders, *Orthographies and reading: Perspectives from Cognitive Psychology, Neuropychology, and Linguistics,* New Jersey: Lawrence Erlbaum Associates.

Coltheart, Max, Patterson, Karalyn E., and Marshall, John. C.

1980 *Deep Dyslexia,* London: Routledge and Kegan Paul.

Ellis, Andrew W. and Young, Andrew W.

1988 *Human Cognitive Neuropsychology,* New Jersey: Lawrence Erlbaum

Associates.

Fang, Sheng-Ping, Horng, Ruey-Yun, and Tzeng, Ovid

1986 Consistency Effects in the Chinese Character and Pseudo-Character Naming Tasks, *Linguistics, Psychology, and the Chinese Language,* Hong Kong: Centre of Asian Studies.

Forster, Kenneth I.

1976 Accessing the Mental Lexicon. *New approaches to the language mechanisms,* Amsterdam: North Holland.

Funnell, Elaine

1983 Phonological Processes in Reading: New Evidence from Acquired Dysgraphia, *Bristish Journal of Psychology,* 74: 159-180.

Glushko, Robert J.

1979 The Organization and Activation of Orthographic Knowledge in Reading Aloud, *Journal of Experimental Psychology: Human Perception and Performance,* 5: 674-691.

Goodman, Roberta Ann and Caramazza, Alfonso

1986 Aspects of the Spelling Process: Evidence from a Case of Acquired Dysgraphia, *Language and Cognitive Processes,* 1. 4: 263-296.

Hoosain, Rumjahn

1991 *Psycholingusitic Implications for Linguistic Relativity: A Case Study of Chinese,* New Jersey: Lawrence Erlbaum Associates.

Hu, Yu-Huan, Qiou, Ying-Guan, and Zhong, Gui-Qing

1990 Crossed Aphasia in Chinese: A Clincial Survey, *Brain and Language,* 39: 347-356.

Huang, Chen Yia

1984 Reading and Writing Disorders in Chinese -- Some Theoretical

Issues. *Psychological Studies of the Chinese Language,* Hong Kong: The Chinese Langauge Society of Hong Kong.

Hue, Ching-Wei and Erickson, James R.

1988 Naming and Lexical Decisions for Chinese Characters. Paper presented at the Midwestern Psychological Association, Chicago, IL.

Kohn, S. E. and Friedman, Rhonda B.

1986 Word-Meaning Deafness: A Phonological-Semantic Dissociation, *Cognitive Neuropsychology,* 3: 291-308.

Li, X. T., Hu, C. Q., Zhu, Y. L., and Sun, B.

1984 Neurolinguistic Analysis of Chinese Alexia and Agraphia. *Psychological Studies of the Chinese Language,* Hong Kong: The Chinese Langauge Society of Hong Kong.

Lyman, R. S., Kwan, S. T., and Chao, W. H.

1938 Left Occipito-Parietal Brain Tumor with Observations on Alexia and Agraphia in Chinese and in English, *The Chinese Medical Journal,* 54. 6: 491-516.

Marshall, John C. and Newcombe, Freda

1973 Patterns of Dyslexia: A Psycholinguistic Approach, *Journal of Psycholinguistic Research,* 2: 175-199.

1980 The Conceptual Status of Deep Dyslexia: An Historical Perspective. *Deep Dyslexia,* London: Routledge and Kegan Paul.

McCloskey, Michael, Goodman-Schulman, Roberta, and Aliminosa, Donna

1990 The Structure of Output Orthographic Representations: Evidence from an Acquired Dysgraphic Patient. Paper presented at the meetings of the Academy of Aphasia, Baltimore, Maryland.

McCloskey, Michael and Caramazza, Alfonso
1988 The Case for Single-Patient Studies, *Cognitive Neuropsychology,* 5: 517-528.

Morton, John and Patterson, Karalyn E.
1980 A New Attempt at an Interpretation, or, an Attempt at a New Interpretation. *Deep Dyslexia,* London: Routledge and Kegan Paul.

Naeser, Margaret A., and Chan, Stephen W.-C.
1980 Case Study of a Chinese Aphasic with the Boston Diagnostic Aphasia Exam., *Neuropsychologia,* 18: 389-410.

Newcombe, Freda and Marshall, John C.
1980 Transcoding and Lexical Stabilization in Deep Dyslexia, *Deep Dyslexia,* London: Routledge and Kegan Paul.

Nolan, Karen A. and Caramazza, Alfonso
1983 An Analysis of Writing in a Case of Deep Dyslexia, *Brain and Language,* 20: 305-328.

Patterson, Karalyn E.
1980 Derivational Errors. *Deep Dyslexia,* London: Routledge and Kegan Paul.

Saffran, Eleanor M., and Marin, Oscar S. M.
1977 Reading without Phonology, *Quarterly Journal of Experimental Psychology,* 29: 515-525.

Saffran, Eleanor M.
1982 Neuropsychological Approaches to the Study of Language, *British Journal of Psychology,* 73: 317-337.

Shallice, Tim, Warrington, Elizabeth K., and McCarthy, Rosaleen
1983 Reading without Semantics, *Quarterly Journal of Experimental*

Psychology, 35A: 111-138.

Shallice, Tim, and Warrington, Elizabeth K.

1975 Word Recognition in a Phonemic Dyslexic Patient, *Quarterly Journal of Experimental Psychology,* 27: 187-199.

1980 Single and Multiple Componentcentral Dyslexic Syndromes. *Deep* Dyslexia, London: Routledge and Kegan Paul.

Shallice, Tim

1988 *From Neuropsychology to Mental Structure,* Cambridge: Cambridge University Press.

Tzeng, Ovid J. L., Hung, Daisy L., Chen, Sylvia, Wu, Jori, and Hsi, Mao-Song

1986 Processing Chinese Logographs by Chinese Brain Damaged Patients. *Graphonomics: Contemporary Research in Handwriting,* North-Holland: Elsevier Science Publishers B.V.

Tzeng, Ovid J. L., Hung, Daisy L., Lu, Ching Ching, Chen, Ming Chen, and Hu, Han Hua

1992 Dissociation between Comprehension and Naming in the Visual Processing of Chinese Characters. Paper presented at the first Asian Conference in Psychology, Singapore. (Organized by the Singapore Psychological Association, October 28-30.)

Warrington, Elizabeth K. and McCarthy, Rosaleen

1983 Category Specific Access Dysphasia, *Brain,* 106: 859-878.

Warrington, Elizabeth K. and Shallice, Tim

1984 Category-Specific Semantic Impairments, *Brain,* 107: 829-854.

Zhu, Y.P.

1987 Analysis of Cueing Functions of the Phonetic in Modern Chinese, Unpublished paper, East China Normal University.