Introduction: new frontiers in Chinese psycholinguistics

Ping Li, Li Hai Tan, Elizabeth Bates, and Ovid J. L. Tzeng

A large body of knowledge has accumulated, especially in the last three decades, on the cognitive processes and brain mechanisms underlying language use, language acquisition, and language disorders. Much of this knowledge has come from studies of Indo-European languages, in particular, English. This is no surprise, given the long tradition of scholarly work in these languages and the linguistic and psycholinguistic theories that are produced therein. Some researchers believe that because of the universal principles of language, theories of language and language processing should apply in the same way to all languages even if they are built on facts from specific languages. This universality perspective, reflected most clearly in Chomsky's theories of language, has dominated much of linguistics and psycholinguistics for the last fifty years. Others, however, think that language-specific variations are sufficiently strong to warrant different conceptualizations of linguistic principles and cognitive underpinnings for different languages. Unlike generative theories of language, this second perspective itself is a mixed bag, from the strongest form of the Sapir–Whorf hypothesis that argues for linguistic determinism to modern-day psycholinguistic theories that emphasize language variation and competition. The tension between these two perspectives has yielded much debate in the cognitive and psycholinguistic studies of language, and it is against this backdrop that we see a surge of research interest in recent years in the study of non-Indo-European languages. Our handbook provides a timely synthesis of the debates emerging out of this research interest, in particular, of the psycholinguistic study of the Chinese language.

Properties of the Chinese language

In this volume, we take a broad definition of "Chinese" and use it to refer to all dialectal variations of the Chinese language, including Mandarin Chinese and Cantonese. Spoken by one-fifth of the world's population, Chinese differs significantly from most Indo-European languages and offers unusual features in its orthographic, phonological, lexical, and syntactic structures. Below we provide a brief sketch of some specific properties of the Chinese language

and, where appropriate, introduce the studies (chapters in this handbook) that are relevant to the properties. Readers interested in the linguistic details should consult Chao (1968), Li and Thompson (1981), and P. Chen (1999) for Mandarin Chinese, and Matthews and Yip (1994) for Cantonese.

First, on the orthographic level, Chinese uses characters rather than alphabetic letters as the basic writing unit, in square configurations that map onto meaningful morphemes rather than phonemes in the spoken language. Thus, a Chinese character has a more direct connection with its meaning than a written word in English does. While the majority of Chinese characters are composed of multiple semantic and phonetic radicals, most of these radicals are also characters in their own right. Processing within this fractal configuration may require a cognitive system that differs in important ways from the cognitive system of English and other alphabetic languages. A significant number of chapters in this handbook deal with how Chinese characters are processed, acquired, and represented in the brain (Chen, Weekes, Peng & Lei, chapter 15; Feng, chapter 16; Honorof & Feldman, chapter 17; Leong, chapter 6; McBride-Chang & Zhong, chapter 7; Liu, Wu, Sue & Chen, chapter 19; Peng & Jiang, chapter 31; Perfetti & Liu, chapter 20; Shu & Wu, chapter 9; Taft, chapter 21; Tan & Siok, chapter 32).

On a phonological level, Chinese uses a tonal system to distinguish lexical items, with each item carrying a particular tone (up to nine tones in some dialects). Such tonal information helps to disambiguate lexical items, although it does not eliminate semantic ambiguity. Several chapters in the handbook are related to the processing and acquisition of tones by native and second language speakers (Gandour, chapter 28; Jongman, Wang, Moore & Sereno, chapter 18; Wang, Sereno & Jongman, chapter 22). On a lexical-morphemic level Chinese has a massive number of homophones: according to the Modern Chinese Dictionary (Institute of Linguistics, 1985), 80 percent of the monosyllables (differentiated by tones) in Chinese are ambiguous between different meanings, and 55 percent have five or more homophones. The single syllable *yi* with the dipping tone has up to ninety homophones. Ambiguity exists not only at the morphemic level, but also at the word level. Linguists have debated intensely on whether Chinese has grammatical categories, due to the lack of inflectional morphology (see discussion below) and the large number of words that are ambiguous between categories (e.g. noun–verb homophones). The processing and the neural representation of such ambiguities are discussed by Zhang, Wu and Yip (chapter 24) and Tan and Siok (chapter 32).

On a grammatical level, first, Chinese does not have devices that indicate differences in tense, number, gender, or case; in other words, grammatical functions and relations for sentence constituents are not linked by morphological associations. The lack of inflectional morphology in Chinese has led to various

perspectives on the status and processing of Chinese words and grammar (Kao, 1990; Li, Bates & MacWhinney, 1993; Li, 1996a; Lu et al., 2000). Chinese does have some grammatical morphemes such as the object marker *ba* and the agent marker *bei*, but these are often optional in natural speech and their functions are volatile. Two particular classes of morphemes that are distinctive of Chinese are the aspect markers for verbs and classifiers for nouns. Aspect markers are used to describe the temporal contours of events (e.g. completed versus ongoing), whereas the classifiers are used to describe objects in terms of shape, orientation, dimension, texture, and animacy. Finally, on a syntactic level, Chinese involves a relatively free word order and a high degree of ellipsis. Given a clear context, a sentence can have null subjects or null objects, sounding telegraphic in a richly inflected language when literally translated. These grammatical and syntactic properties and their representation and processing consequences are discussed in the chapters by Cheung and Clark (chapter 1), Erbaugh (chapter 3), Fletcher, Stokes and Wong (chapter 27), Huang (chapter 4), Jia (chapter 5), and Packard (chapter 30).

It is important to emphasize, however, that the language-specific properties in Modern Chinese are the natural outcome of language evolution, and should not be viewed mystically or suspiciously. A careful study of Chinese and its psycholinguistic aspects should thus help us to dispel the myths and mysteries about this language that have surfaced again and again in the media and even in scholarly work, as both Au (chapter 25) and Erbaugh (chapter 3) point out.

Overview of theoretical issues

The specific properties of Chinese raise interesting questions about the mechanisms of language acquisition, language processing, and language and the brain. One the one hand, Chinese presents interesting challenges to current psycholinguistic theories, and on the other, it also presents new windows on cognitive processes and new opportunities for psycholinguistic investigation. In the past three decades, researchers interested in Chinese have employed a variety of theoretical and experimental paradigms to study language acquisition and processing, and in each case the language-specific properties of Chinese have served as crucial test cases for general theories or models of psycholinguistics. Below we briefly summarize the current state of the art with five general issues or directions that researchers have focused on in the past decades. These are not meant to be exhaustive in any way, but are discussed here to highlight the significance of some recurring issues.

1. *Role of orthography versus phonology in lexical (semantic) processing*. In the past two decades a large body of research in Chinese psycholinguistics has investigated how orthographic and phonological information contributes to Chinese character identification. A fierce debate in the literature is whether

the reading of Chinese involves automatic phonological access, and if so, how early phonological codes are activated, relative to the orthographic features of characters and semantic information of words (Perfetti & Tan, 1998; Pollatsek, Tan & Rayner, 2000; Tan, Hoosain & Peng, 1995; Tan, Hoosain & Siok, 1996; Weekes, Chen & Lin, 1998). While a number of studies demonstrated semantic activation with the early involvement of phonology (Chua, 1999; Perfetti & Zhang, 1995; Spinks et al., 2000; Tan & Perfetti, 1997; Xu, Pollatsek & Potter, 1999), other studies did not (Chen, Flores d'Arcais & Cheung, 1995; Chen & Shu, 2001; Zhou & Marslen-Wilson, 1999). More recently, researchers have been concerned with the question of what is the nature of phonological and orthographic codes in Chinese (Chen et al., this volume, chapter 15; Feldman & Siok, 1999a, b; Liu et al., this volume, chapter 19; Spinks et al., 2000; Taft, this volume, chapter 21; Tan & Siok, this volume, chapter 32; Xu, Pollatsek & Potter, 1999; Zhang, Perfetti & Yang, 1999). Relevant to the "nature" issue is how reading development and phonological knowledge interact in the process of children's acquisition of Chinese characters. Recent studies indicate that children's knowledge and awareness of the phonological structure of characters contributes significantly to their reading abilities (e.g. Leong, chapter 6; McBride-Chang & Zhong, chapter 7; Shu & Wu, chapter 9, all in this volume, and Siok & Fletcher, 2001).

2. *Time course of lexical access in sentence processing*. Apparently, the time course of the activation and utilization of linguistic components is central to any theory of language processing, as seen in the debates on phonological and semantic activation in character recognition. In auditory sentence processing, researchers have been concerned with how meanings of individual words are accessed during the course of comprehension. The past twenty-five years have seen numerous studies in this domain, but the major theories of language processing derived from these studies differ roughly along one dimension: whether lexical access is completed first to await the fitting to a sentence context (modular) or whether context can provide an early top-down mechanism (interactive) during the access stage. Chinese provides a unique test case in this regard, because of the extensive level at which word forms are ambiguous. Research conducted by investigators has looked at spoken word recognition with respect to the time course of lexical access, in the bilingual context (Li, 1996b; Li & Yip, 1998) as well as the monolingual and aphasia contexts (Li & Yip, 1998; Lu et al., 2002; Wu & Shu, 2002; Zhang et al., 2001; Zhang, Wu & Yip, this volume, chapter 24). An exciting development for the understanding of this issue is the use of neuroscience methods, including event-related-potentials (ERP) and functional magnetic resonance imaging (fMRI), as discussed in Zhang, Wu and Yip (this volume, chapter 24).

3. *Interaction of lexicon, grammar, and context in acquisition*. Chinese-speaking children display very early sensitivity to language-specific characteristics, including the early acquisition of aspect markers, the *ba* and *bei* constructions, and the lexical compounding process (see Erbaugh, 1982, 1992; Lee, 1996). The interaction between the acquisition of grammatical morphemes, the semantics of verbs, and the contextual cues is most clearly reflected in children's acquisition of the aspect markers (e.g. *-le, zai*, *-zhe*; see Li & Bowerman, 1998; Li & Shirai, 2000). In addition, the high degree of lexical encoding of syntactic information in verbs, coupled with the important role of contextual interpretation, provides the child with opportunities to explore various form-to-form, form-to-meaning, and meaning-to-context relationships in the input speech by parents or caregivers (this volume: Cheung & Clark, chapter 1; Huang, chapter 4; Fletchers, Stokes & Wong, chapter 27; Shi, chapter 8; Li, chapter 29). Thus, Chinese provides an ideal test ground for theories of both conceptual and linguistic development. Recent research in this domain indicates that in contrast to English and other Indo-European languages, verbs emerge earlier and in larger quantity than nouns for Mandarin-speaking children (Tardif, 1996, this volume, chapter 19; Tardif, Shatz & Naigles, 1997). However, nominal classifiers provide a situation counter to this earliness in acquisition: early on children often confuse the use of classifiers, demonstrating the interaction between a probabilistic learning mechanism and the complex feature binding process embedded in categorization and lexical classification (Erbaugh, this volume, chapter 3).
4. *Neuroanatomical mechanisms of processing and acquisition*. Because written Chinese differs notably from alphabets in visual form, and spoken Chinese differs in syntactic and semantic structure from western languages, the uncovering of neuroanatomical mechanisms in the processing and acquisition of Chinese has now become increasingly important to the understanding of the universality or the variation of brain organization of language. Most of the past research in this domain has adopted the visual hemi-field paradigm in an attempt to ascertain whether a Chinese character's unique square configuration leads to the processing dominance in the right cerebral hemisphere. So far, the conclusions have been contentious (Tzeng et al., 1979; Besner, Daniel & Slade, 1982; Cheng & Yang, 1989; Fang, 1997). More recent investigations with neuroimaging techniques such as fMRI suggest that there is no interhemispheric difference between the processing of Chinese and the processing of English (Chee, Tan & Thiel, 1999; Chee et al., 1999; Tan & Siok, this volume, chapter 32), and that hemispheric lateralization for Chinese-speaking children is accomplished by about 7 years of age (Kwok et al., 2003). Gandour (this volume, chapter 28; Gandour et al. 2003c) also showed that in the perception of speech prosody, functionally

relevant properties of the complex auditory stimuli are critical in determining hemispheric lateralization. However, reading in Chinese does result in some important neural activity (in neural circuits involving the left middle frontal cortex and the right hemisphere) that is rarely observed during the reading of English, according to Tan et al. (2000, 2001). New fMRI data also suggest that the neural representation of grammatical categories might differ significantly between Chinese and Indo-European languages, in that Chinese nouns and verbs do not evoke distinct cortical responses as do their counterparts in English and other languages (Li, Jin & Tan, 2004).

5. *Neural network modeling and computational analyses*. Although neural network modeling and computational analyses have become extremely influential and popular in the study of Indo-European languages, the application of these methods to the psycholinguistic study of Chinese is very new. There have been only a few published works using neural networks to model or to explain word recognition, sentence processing, and language acquisition in Chinese. Taft (1994; this volume, chapter 21) suggested an approach to character recognition on the basis of the classic interactive-activation model of word recognition of McClelland and Rumelhart (1981). Chen and Peng (1994; see also Peng & Jiang, chapter 31) and Perfetti and Liu (this volume, chapter 20) used multi-layered perceptrons with back-propagation learning to model the interactive activation of single character recognition. Li (2002b; this volume, chapter 29) and Xing, Shu, and Li (2002) applied self-organizing neural network models to the study of the acquisition of Chinese characters, word categories, and bilingual lexical processing, illustrating how the acquisition of words and characters via statistical characteristics of the input can yield systematic patterns in representation that match empirical data. Because of the important role that computational approaches have played in identifying mechanisms of processing and acquisition, psycholinguists can no longer afford to disregard the development of computational models that allow researchers to study complex issues in Chinese that would be difficult to address in traditional ways (see Erbaugh, chapter 3, Hernandez, Li & MacWhinney, 2005, and Stokes, chapter 10, this volume, for examples and discussion).

Each of these five topics or directions has been discussed in some depth in the chapters of the handbook. As many of these issues are highly related, our handbook provides a context where fruitful bridging of different perspectives is possible.

Scope and structure of the handbook

Despite the significant impact of psycholinguistic research in Chinese and other East Asian languages, so far there have been relatively few efforts to integrate the various perspectives in this vast amount of literature. Our handbook aims

to fill this gap. Initially we had in mind two groups of readers as the primary audience of our handbook: psycholinguists who are interested in crosslinguistic and comparative studies of Indo-European and East Asian languages, and graduate students who are interested in doing research in the fast-expanding field of East Asian psycholinguistics. As it turns out, our handbook is suitable not only for the psycholinguists and graduate students, but also for a wide range of other readers with an international background, including cognitive psychologists, linguists, neuroscientists, educators, and other professionals interested in language. In each of the thirty-two chapters, the authors have attempted to provide a state-of-the-art overview, summarizing key issues and debates in the relevant field or subfield and providing approaches to tackle these issues within and across their domains of expertise. The various topics of the chapters cut across traditional disciplinary boundaries, and cover a wide variety of issues in language, cognition, culture, and brain. In addition to its academic and scientific value, readers may also find that many of the discussions have practical implications for language education, including the learning and teaching of Chinese as a native (e.g. the role of epilinguistic awareness in reading, as shown by Leong, chapter 6) and as a foreign language (e.g. the tone perception and training studies, as reported in Wang, Sereno & Jongman, chapter 22).

Chapters in the handbook provide a roadmap to topics within three general domains of psycholinguistic research: language acquisition, language processing, and language and the brain. First, there are thirteen chapters dedicated to language acquisition, and the topics include children's acquisition of verbs and temporal references (Cheung & Clark, chapter 1; Huang, chapter 4), classifiers (Erbaugh, chapter 3), and grammatical and syntactic categories (Shi, chapter 8; Tardif, chapter 11); reading acquisition in Chinese (Leong, chapter 6; McBride-Chang & Zhong, chapter 7; Shu & Wu, chapter 9); acquisition of binding principles and parameter setting (Chien & Lust, chapter 2; Yang, chapter 12); acquisition of phonology (Stokes, chapter 10); and second language or bilingual acquisition (Jia, chapter 5 ; Yip, chapter 13). For language processing, we have eleven chapters, with a large amount dedicated to the processing of Chinese characters (Chen et al., chapter 15; Feng, chapter 16; Honorof & Feldman, chapter 17; Liu al., chapter 19; Perfetti & Liu, chapter 20; Taft, chapter 21); other chapters in this category include topics on speech production (Chen & Dell, chapter 14), speech perception and processing (Jongman et al., chapter 18; Wang, Sereno & Jongman, chapter 22), lexical ambiguity processing (Zhang, Wu & Yip, chapter 24), and discourse processing (Yang, Gordon & Hendrik, chapter 23). Finally, for topics on language and the brain we have eight chapters, including language impairment in children (Fletcher, Stokes & Wong, chapter 27) and in patients (Packard, chapter 30), functional neuroimaging of bilingual processing (Chee, chapter 26), neuorimaging of Chinese speech prosody (Gandour, chapter 28), of Chinese character processing (Peng & Jiang, chapter 31), and of Chinese reading (Tan & Siok,

chapter 32), neural network modeling (Li, chapter 29), and language and cognition (Au, chapter 22). Note that the classification of the thirty-two chapters into the three general domains (sections) may seem arbitrary at times: all chapters cut across boundaries to some extent, and some chapters clearly deal with both acquisition and processing, and some even with all three domains. Although the general headings under which the chapters are listed may not accurately reflect the contents of the discussion, they are nevertheless useful in serving as conceptual breaks when readers sift through the nearly three dozen papers.

While we recognize that it is impossible to represent all exciting research in the field in one volume, we believe that our handbook provides a balanced view on most of the important issues currently under debate. Contributions to our volume include studies that are based on a variety of conceptual, theoretical, and methodological foundations, from cognitive approaches to linguistic approaches, from formal theories to functional theories, from classic psycholinguistic paradigms to computational models, and from behavioral studies to neuroimaging techniques. The topics of the chapters are diverse as well as comprehensive. In some cases there is not the perfect balance, which often reflects the status quo of the field. For example, we have fewer chapters in the language and brain section, as compared with the language acquisition section, reflecting the fact that research in the latter for Chinese is in its early stages; on the other hand, the weight clearly tilts toward Chinese characters in the language processing section, due to the unique features of the Chinese writing system and the large amount of research therein.

Returning to the starting point of this introduction, we are still far from reaching a conclusion on the debate regarding language universals and language variations. Research summarized in our handbook clearly gives evidence of both. While some studies emphasize universality and innateness for language acquisition (Chien & Lust, chapter 2; Yang, chapter 12), others stress the importance of differences between Chinese and other languages (Fletcher, Stokes & Wong, chapter 27; Tardif, chapter 11; Zhang, Wu & Yip, chapter 24). While some argue for the significant processing and neural consequences of the language-specific properties of Chinese (Gandour, chapter 28; Tan & Siok, chapter 32), others opt for language-neutral principles that characterize acquisition, processing, and breakdown (Feng, chapter 16; Packard, chapter 30; Yang, Gordon & Hendrick, chapter 23). The opposition between these competing perspectives reflects an even deeper contrast of opinions on the relationship between language, culture, and cognition (Au, chapter 25; Erbaugh, chapter 30; Tardif, chapter 11), and that brings us back to the debate on the Sapir–Whorf hypothesis and various other contentious issues about how the mind and brain work in different languages and cultures. The psycholinguistic study of Chinese alone cannot resolve these issues, but it is our genuine hope that this handbook serves as a catalyst for

new research on the cognitive and neural mechanisms underlying language in general.

Acknowledgments

Preparation of this article was supported by grants from the National Science Foundation (BCS-0131829) and the Hong Kong Research Grants Council (HKU 7133/01H). Partial support was also provided by visiting fellowships from the University of Hong Kong while the first author was on leave in the Division of Speech and Hearing Sciences and the Joint Laboratories for Language and Cognitive Neuroscience (recently named as the State Key Laboratory of Brain and Cognitive Sciences), University of Hong Kong.

Part I

Language acquisition