A psycholinguistic database for traditional Chinese character naming

Ya-Ning Chang · Chun-Hsien Hsu · Jie-Li Tsai ·
Chien-Liang Chen · Chia-Ying Lee

Abstract In this study, we aimed to provide a large-scale set of psycholinguistic norms for 3,314 traditional Chinese characters, along with their naming reaction times (RTs), collected from 140 Chinese speakers. The lexical and semantic variables in the database include frequency, regularity, familiarity, consistency, number of strokes, homophone density, semantic ambiguity rating, phonetic combinability, semantic combinability, and the number of disyllabic compound words formed by a character. Multiple regression analyses were conducted to examine the predictive powers of these variables for the naming RTs. The results demonstrated that these variables could account for a significant portion of variance (55.8 %) in the naming RTs. An additional multiple regression analysis was conducted to demonstrate the effects of consistency and character frequency. Overall, the regression results were consistent with the findings of previous studies on Chinese character naming. This database should be useful for research into Chinese language processing, Chinese education, or cross-linguistic comparisons. The database can be accessed via an online inquiry system (http://ball.ling.sinica.edu.tw/namingdatabase/index.html).

Y.-N. Chang · C.-H. Hsu · C.-L. Chen · C.-Y. Lee (✉)
Institute of Linguistics, Academia Sinica, Taipei City, Taiwan
e-mail: chiaying@gate.sinica.edu.tw

J.-L. Tsai
Department of Psychology, National Chengchi University, Taipei
City, Taiwan

C.-Y. Lee
Institute of Neuroscience, National Yang-Ming University,
Taipei City, Taiwan

C.-Y. Lee
Institute of Cognitive Neuroscience, National Central University,
Taoyuan City, Taiwan

J.-L. Tsai
Research Center for Mind, Brain and Learning, National Chengchi
University, Taipei City, Taiwan

Understanding how readers process words, which are fundamental units for reading, has been a core issue in psycholinguistic studies. Behavioral studies have utilized tasks such as naming, lexical decision, speech perception, and synonym judgment, with the aim of understanding how orthographic, phonological, and semantic information is processed during visual word recognition. These experiments have typically employed a factorial design by manipulating one or two lexical variables (e.g., word frequency and spelling-to-sound consistency) in a limited set of stimuli, while trying to control for other lexical variables (e.g., orthographic neighborhood size and acoustic properties of onsets) in order to prevent any potential confounds. Although this approach has provided fruitful evidence to advance our understanding of how various lexical properties affect visual word recognition, it has some limitations in terms of how to select a set of items that vary in only one dimension and how to generalize the findings to a larger set of stimuli.

To overcome these limitations, a number of large-scale psycholinguistic databases have been made available for languages with alphabetic orthographies, including English (Balota et al., 2007; Coltheart, 1981; Keuleers, Lacey, Rastle, & Brysbaert, 2012), French (Ferrand et al., 2010), Dutch (Keuleers, Diependaele, & Brysbaert, 2010), Italian (Barca, Burani, & Arduino, 2002), and Malay (Yap, Rickard Liow, Jalil, & Faizal, 2010). For example, in English, the MRC Database provides a set of word association norms, as well as syntactic information (Coltheart, 1981). Furthermore,

the English Lexicon Project (ELP) has collected large-scale normative data, including descriptive characteristics for over 40,000 words and nonwords, along with their average reaction times (RTs) from over 1,200 participants (Balota et al., 2007). This database has been used in several regression analyses of naming and lexical decision tasks (Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004; New, Ferrand, Pallier, & Brysbaert, 2006) and for evaluations of computational models (Perry, Ziegler, & Zorzi, 2007). Recently, the British Lexicon Project (BLP), developed by Keuleers et al. (2012), has provided individual RTs for 28,739 English words. These standardized psycholinguistic norms not only provide a window to investigate how various lexical properties simultaneously affect visual word recognition, but also greatly reduce the time and effort spent preparing stimuli.

Regarding Chinese, only recently have large-scale psycholinguistic norms for simplified Chinese characters been developed (Liu, Shu, & Li, 2007; Sze, Rickard Liow, & Yap, 2013). Liu et al. provided the first large-scale database of simplified Chinese character naming. This database of 2,423 simplified Chinese characters includes 15 lexical–semantic variables (e.g., age of acquisition, character frequency, familiarity, concreteness, and imageability) that are important for predicting naming RTs. Liu et al. demonstrated that age of acquisition and character frequency had the greatest effects on naming. Semantic variables also made unique contributions. Although they provided an online resource for the lexical variables, the available naming data were not included. More recently, Sze et al. collected and published RTs for a set of 2,500 simplified characters in a lexical decision task. These data are potentially valuable as a resource for research that uses either a factorial design or a regression approach. Nevertheless, the lexical properties of those characters were not included in their study.

However, two different Chinese writing systems—simplified and traditional scripts—are used in different Chinese-speaking regions. The simplified characters are commonly used in Mainland China, Singapore, and Malaysia, whereas the traditional characters are currently used in Taiwan and Hong Kong. Simplified characters are created from the traditional characters by using a series of simplification processes: for example, by simplifying character radicals (e.g., 請→请, "qing3"), by reducing the number of character strokes (e.g., 華→华, "hua2"), or by replacing a character with another existing character (e.g., (穀, 谷)→谷, "gu3") (McBride-Chang, Chow, Zhong, Burgess, & Hayward, 2005). This results in generally fewer strokes and homophones in simplified than in traditional characters. It is worth noting that the two existing sets of Chinese norms (Liu et al., 2007; Sze et al., 2013) are based on simplified Chinese characters. The existing norms may not be directly applied in studies on traditional characters, in particular for word-form and frequency-related measures. In this study, we thus aimed to create the first large-scale database that provides ten lexical variables

and naming RTs for 3,314 traditional Chinese characters, for use in future research.

Studies on Chinese character processing have demonstrated that a number of lexical–semantic variables affect naming RTs (Fang, Horng, & Tzeng, 1986; Feldman & Siok, 1999; Lee, Tsai, Su, Tzeng, & Hung, 2005; Li, Bi, Wei, & Chen, 2011). Variables such as frequency, regularity, age of acquisition, imageability, familiarity, and concreteness have been included in the existing psycholinguistic norms (Liu et al., 2007). Nevertheless, some variables have not yet been considered, in particular those related to radical information. For example, one important yet overlooked variable is orthography-to-phonology mapping consistency (Glushko, 1979; Plaut, McClelland, Seidenberg, & Patterson, 1996; Seidenberg & McClelland, 1989; Taraban & McClelland, 1987). In alphabetic languages like English, the degree of spelling-to-sound consistency can be measured at the rime level. Words (e.g., WAVE–CAVE) that share an orthographic body and are pronounced in the same way are classified as friends; otherwise, they are enemies (e.g., WAVE–HAVE). The degree of spelling-to-sound consistency can be measured by types (the total number of friends to the total number of neighbors) or by tokens (taking the frequency of each orthographic neighbor into account) (Jared, 1997). In general, low-consistency words take longer to pronounce than high-consistency words (Glushko, 1979; Jared, 1997; Jared, McRae, & Seidenberg, 1990).

Chinese has a logographic writing system with deeper orthography than alphabetic languages—that is, there is no simple mapping between orthography and phonology. However, approximately 80 % of traditional Chinese characters are phonograms, consisting of a semantic radical and a phonetic radical (Zhou, 1978). The semantic radical (usually on the left) carries information regarding meaning, while the phonetic radical (usually on the right) provides a clue regarding the character's pronunciation. For those phonograms, the orthography-to-phonology consistency of a given character indicates whether its pronunciation agrees with that of other characters containing the same phonetic radical. The consistency score, defined by type, is a measure of the total number of friends to the total number of neighbors. Similar to the findings in English, several studies have shown consistency effects in naming Chinese characters (Fang et al., 1986; Hue, 1992; Yang, McCandliss, Shu, & Zevin, 2009). However, these studies have been inconsistent as to whether the consistency effect emerged in naming high-frequency characters. To resolve this, Lee et al. (2005) manipulated the consistency measure by token (i.e., by taking character frequencies into consideration) in a naming study. They demonstrated consistency effects for both high- and low-frequency characters, and a significant interaction between consistency and frequency, in which the consistency effect was more prominent in low-frequency than in high-frequency words. They also suggested

that controlling the number and the frequency of friends and enemies might be the key to showing consistency effects for high-frequency characters.

Another important lexical factor in Chinese reading is radical combinability, also called *radical frequency by type*, which is defined as the number of phonograms that share the same radical (Feldman & Siok, 1999). Several studies have shown that radical combinability has an effect on response latencies in Chinese character processing (Chen & Weekes, 2004; Feldman & Siok, 1999; Li et al., 2011). Radical combinability can be subdivided into phonetic combinability (the number of phonograms sharing the same phonetic radical) and semantic combinability (the number of phonograms sharing the same semantic radical) by radical function. Facilitative combinability effects for both phonetic and semantic radicals have been found in a character decision task (Feldman & Siok, 1999). Moreover, characters with greater semantic combinability are processed more quickly in a semantic category task than are characters with lower semantic combinability (Chen & Weekes, 2004). However, Li et al. demonstrated an inhibitory effect of phonetic combinability when naming inconsistent characters, but reported a null effect when naming consistent characters. It seems that whether radical combinability facilitates or inhibits processing depends on different task demands. Nevertheless, all of these findings show that both phonetic and semantic radicals are important units in processing Chinese characters.

In addition to single-character studies, there is also evidence that the lexical processing of Chinese disyllabic compound words is affected by the words' constituent characters. Specifically, the more neighbors that share the same constituent character with a disyllabic compound word, the greater the response latency is for identifying that word (Huang et al., 2006; Tsai, Lee, Lin, Tzeng, & Hung, 2006), particularly when the constituent character is embedded in the first position of the word. It is unclear whether the number of times that a character is used in compound formation also plays a role in predicting character naming RTs. The subjective semantic ambiguity for these characters is also included in the present database, since they may affect the reading of Chinese words (Hsu, Lee, & Marantz, 2011; Huang, Lee, Huang, & Chou, 2011).

In sum, in the present study we describe an attempt to provide psycholinguistic norms for 3,314 traditional Chinese characters, to complement existing research norms. This database includes five variables (i.e., consistency, phonetic combinability, semantic combinability, semantic ambiguity rating, and number of disyllabic compound words formed by a character) that have not been considered previously, and five variables (i.e., frequency, regularity, homophone density, familiarity, and number of strokes) that have been shown to be important predictors for naming RTs (Liu et al., 2007). In total, ten variables are included in the present database. In

addition, behavioral naming RTs are provided to substantiate the variables. To further assess the database and demonstrate the importance of selecting ranges of lexical variables, additional multiple regression analyses were conducted to explore the effects of consistency and character frequency (Lee et al., 2005; Yang et al., 2009). The resulting database will be useful for researchers in the area of Chinese language processing, Chinese educators, or linguists interested in conducting cross-linguistic comparisons. To make these valuable data accessible to researchers and educators, an online inquiry system has been developed (http://ball.ling.sinica.edu.tw/namingdatabase/index.html).

Method

A set of 3,314 traditional Chinese characters were included in the present database. These consisted of phonograms and nonhomographs, selected from amongst 5,640 characters in the Academia Sinica Balanced Corpus (ASBC; Huang & Chen, 1998). The ASBC corpus consists of 9,227 articles (approximately five million words) covering the fields of philosophy, science, social society, art, lifestyle, and literature. Most lexical variables included in the present database were also derived from the ASBC; others, such as familiarity and semantic ambiguity, were collected from subjective ratings or taken from our previous work. The details of all lexical variables and naming RTs are described as follows.

Lexical variables

Character frequency A considerable amount of evidence has indicated that character frequency is a powerful predictor of RTs in Chinese character recognition; frequency effects have been observed consistently in a wide range of tasks (Hue, 1992; Lee et al., 2005; Liu et al., 2007; Yang et al., 2009). Higher frequency characters tend to be processed more quickly than lower frequency characters during naming (Lee et al., 2005; Yang et al., 2009). In the present study, the character frequency data was compiled from the ASBC (Huang & Chen, 1998). The corpus contains approximately five million words, and the character frequency is defined as the number of times that a character appears per million. The frequency measure in the database ranged from 1 to 60,158.

Regularity *Regularity* is defined as whether a character is pronounced the same as its phonetic radical (Fang et al., 1986). The regularity of characters used here was categorized into three groups, represented by numerical values: 0 indicates *irregular*, 1 indicates *regular*, and 2 indicates that the phonetic radical is *unpronounceable*. For example, 佳 "jia1" would be rated as 0, 潭 "tan2" as 1, and 流 "liu2" as 2.

Consistency As in studies in English (Jared et al., 1990), consistency can be measured in two different ways. The first consistency value defined by type is the ratio of the number of characters sharing a phonetic radical that have the same pronunciation, to the total number of characters sharing that phonetic radical (Fang et al., 1986). The second consistency measure defined by token takes character frequency into consideration and is the ratio of the summed frequencies of characters sharing a phonetic radical that have the same pronunciation, to the summed frequencies of characters sharing that phonetic radical (Lee et al., 2005; Lee, Huang, Kuo, Tsai, & Tzeng, 2010). In the present database, both measures were derived from the ASBC (Huang & Chen, 1998). Tonal differences were discarded, and character frequency was logarithmically transformed. Both scores ranged from 0 to 1. For example, eight characters share the same phonetic radical 登 ("deng1"). Only 橙 and 澄 are pronounced in the same way (i.e., "cheng2") so they have a consistency value of 0.25 (i.e., 2/8). When character frequencies are taken into account, the consistency value of the characters 橙 and 澄 is 0.243. For polyphonic characters (i.e., characters that have more than one pronunciation), their consistency values were computed according how frequently their pronunciations are used (Ministry of Education, Taiwan, 2012). For example, the character 估 can be pronounced as "gu1" or "gu4," but the dominant pronunciation is "gu1," and the pronunciation "gu4" is barely used. Thus, we removed the infrequent one from the computation. Another example, 說, has two pronunciations (i.e., "shou1" and "shui4") that are used equally frequently. They were treated as two independent characters and were included in the computation. The same rules were applied for characters with more than two pronunciations.

Phonetic combinability and semantic combinability According to Feldman and Siok (1999), phonetic combinability is the number of characters that share the same phonetic radical, also known as *phonetic radical frequency by type*. Similarly, semantic combinability is defined as the number of characters that share the same semantic radical, also known as *semantic radical frequency by type*. The metrics of both phonetic combinability and semantic combinability were based on the norms reported by Hsu et al. (2011), which were derived from 3,967 characters. It is worth noting that the function of radicals is of interest here so the positions of the radicals were not taken into account. The scales of phonetic combinability and semantic combinability ranged from 1 to 20 and from 1 to 226, respectively.

Homophone density In alphabetic systems, homophones are words that have the same pronunciation but differ in spelling and meaning. In Chinese, each character corresponds to a single syllable; homophones in Chinese are thus defined as characters that have the same syllable and tone but different spellings and meanings. Homophone density, then, refers to the number of characters that share the same syllable. For example, five characters (碼, 瑪, 馬, 螞, 嗎) share the same syllable and tone "ma3." On the basis of our analysis of the ASBC (Huang & Chen, 1998), there are approximately 1,200 unique syllables with tones in Chinese. Over 80 % of these are represented by more than one character. The score of homophone density in the present database varied from 1 to 69.

Number of strokes Chinese characters are often described by the number of strokes required to write them. The number of strokes that a character has may indicate the level of its visual complexity. In the present database, the number of strokes ranged from 4 (e.g., 仍 "reng1") to 30 (e.g., 鸝 "li1").

Number of disyllabic compound words formed by a character In Chinese, some characters can be used in free-standing form whereas others are usually combined to form compound words. Since disyllabic compound words are the most common constructions, this metric was computed on the basis of approximately 67,000 disyllabic compound words in the ASBC (Huang & Chen, 1998). In addition, the constituent character embedded in a disyllabic compound word could either lie in the first or second position; therefore, both scores were included in the database. The number of disyllabic compound words formed by a particular character in the first position and in the second position ranged from zero to 119 and from zero to 126, respectively. Both scores were based on word types.

Semantic ambiguity rating A Chinese character is considered semantically ambiguous if multiple meanings are associated with it (Perfetti & Tan, 1998). For example, the character 花 in 花園 'garden' means "flower," whereas the character 花 in 花費 'expense' means "spend." It is therefore difficult to access the meaning of the character 花 out of context. Several studies have used subjective ratings to measure the semantic ambiguity of a character (Hsu et al., 2011; Huang et al., 2011; Perfetti & Tan, 1998; Tan, Hoosain, & Siok, 1996). This could reflect how precisely the meaning of a character is organized in one's mental representation. In the present study, semantic ambiguity scores were based on subjective ratings taken from Huang et al. (2011) and Hsu et al. (2011). In these studies, participants were asked to rate the diversity of meanings when the character was used in a word. The values ranged from 1 to 5, where a score of 1 meant the character only has *one meaning*, whereas a score of 5 meant the character has *highly diverse meanings*.

Familiarity Familiarity refers to the measure of how frequently participants have encountered a character. Previous studies have used subjective ratings as measurements for familiarity

of characters (Balota, Pilotti, & Cortese, 2001; Liu et al., 2007), and this method was adopted here. A total of 5,640 characters from the ASBC were randomly separated into four subsets, and each subset included 1,410 characters. In all, 136 college students participated in the rating. The participants had a mean age of 22 years, all were native Chinese speakers, and none had any history of dyslexia. Each participant was assigned to one subset of stimuli, so each character had a total of 34 ratings. Characters were presented one by one in a random order on the center of a screen; each character remained on the screen until the participant gave a response. Participants were asked to give each character a value from a 7-point scale, ranging from *never encountered* to *encountered several times a day* (i.e., 1 = *never*, 2 = *once a year*, 3 = *once a month*, 4 = *once a week*, 5 = *once every two days*, 6 = *once a day*, 7 = *several times a day*).

Participants

A total of 140 college students with normal (or corrected-to-normal) vision participated in the naming task. Their ages ranged from 18 to 26 years old. All were native Chinese speakers, and none were reported to have any cognitive problems. All participants were paid for their participation in the experiment.

Procedure

Following the procedures used in Balota et al. (2007) and Ferrand et al. (2010), the 3,314 characters were randomly split into seven sublists: six of the sublists consisted of 473 characters, and one sublist consisted of 476 characters. To obtain 20 responses per item, 140 participants were tested (7 × 20 = 140). The presentation order of the stimuli for each participant was randomized. Participants were tested one at a time in a small room. They sat in front of a computer at a distance of approximately 60 cm from the screen. Each trial began with a visual presentation of a fixation point for 600 ms, accompanied by a 200-Hz beep signal for 200 ms, followed by a target character. Participants were asked to read aloud all items as quickly and as accurately as possible. All items were presented in Biaukai 20-point white font in the center of a black screen. The target character remained on the screen until the participant had responded or until an interval of 3, 000 ms had elapsed. Participants were given 20 practice trials to familiarize themselves with the procedure before the experiment began. There were six breaks during the experimental session. Naming RTs were recorded using a voice-key delay. Naming accuracy was hand-coded by the experimenter, and any vague responses were resolved by listening to the recorder.

Analyses

Multiple regression analyses were used (1) to examine whether the lexical variables listed above were significant predictors for naming RTs and (2) to demonstrate the potential use of the database by exploring the relationship between consistency and level of character frequency. Incorrect responses in the naming task were excluded from the analysis. After removing the incorrect responses, the mean and standard deviation were calculated at the item level. Any response three standard deviations above or below the mean was discarded. In total, 1.4 % of the data were discarded, and this resulted in the use of 3,267 characters for the regression analysis.

In this study, naming RTs were collected from different groups of participants who had received different sets of characters. This could potentially have produced large between-subjects variability, although a yoked control procedure was adopted, with the yoked group consisting of seven participants. To alleviate this issue, the participants' naming RTs were standardized using the *z*-transform prior to analysis. Since the distribution of character frequency was strongly skewed, character frequency was logarithmically transformed to fulfill the assumption of normality (Hair, Black, Babin, & Anderson, 2009). In this way, the relationship between frequency and RTs also became linear (Murray & Forster, 2004). Other variables, including homophone density, semantic ambiguity rating, and number of disyllabic compound words, were also logarithmically transformed for the same reasons prior to the regression analyses. In addition, since regularity was a categorical variable that had three levels (i.e., 0 = *irregular*, 1 = *regular*, and 2 = *unpronounceable*), it needed to be dummy-coded for the regression analyses (Hair et al., 2009). Hence, the categorical variable, regularity, was converted into two binary variables, one called *regular* and the other called *unpronounceable*. If the regularity was 0, both *regular* and *unpronounceable* were coded with a 0; if regularity was 1, *regular* was coded with a 1 and *unpronounceable* was coded with a 0; if regularity was 2, *regular* was coded with a 0 and *unpronounceable* was coded with a 1.

To control for potential onset effects caused by the sensitivity of the voice key to different onsets (Balota et al., 2004; Liu et al. 2007) in naming, the initial phonemes were also considered in the following regression analyses. The procedure for coding onsets was adopted from Balota et al. (2004). The initial phoneme of each character was coded dichotomously (1 or 0) for the following 13 features, where 1 denoted the presence of the feature and 0 denoted its absence: stop, affricate, fricative, nasal, liquid, aspirated, voiced, bilabial, labiodental, alveolar, palato-alveolar, alveo-palatal, and velar.

Results

Descriptive statistics

Z-scored RTs were used as a dependent variable in the regression model. The following were included as predictors in the regression model: initial phonemes (IP), character frequency (Freq), number of strokes (NS), regular (REG), unpronounceable (UNP), phonetic combinability (PhonN), semantic combinability (SemN), homophone density (HD), semantic ambiguity rating (SAR), familiarity (FAM), number of disyllabic compound words formed by a character in the first position of the word (NC1), and number of disyllabic compound words formed by a character in the second position of the word (NC2). Regarding the consistency variable, Lee et al. (2005) suggested that naming RTs are affected by both the number and frequency of neighbors. Therefore, consistency with respect to character frequency (CON) was used in the regression model.

Table 1 shows descriptive statistics of all of the predictors and naming RTs. Table 2 shows the intercorrelations between all variables; simple correlations between the naming RTs and the predictors are also included.

As can be seen in Table 2, character frequency is correlated with semantic ambiguity rating, familiarity, and the number of disyllabic compound words formed by a character, indicating that high-frequency characters are judged to be familiar, tend to have multiple meanings, and are often used in compound formations. The consistency score is negatively correlated with phonetic combinability, showing that large phonetic groups tend to be less consistent.

Predictive powers of lexical variables

As is shown in Table 3, Block 1, the initial phonemes were entered in the regression model, and all other variables were then simultaneously entered in Block 2, to examine which of these variables could account for a significant proportion of the variance in z-scored RTs. The results showed that the initial phonemes accounted for a significant portion (adjusted $R^2 = .28$) of the z-scored RTs. When the onset effect was controlled for, Freq, CON, REG, UNP, SAR, FAM, and NC1 all made significant contributions to account for the variance of z-scored RTs, whereas NS, PhonN, SemN, HD, and NC2 did not exert significant effects. The regression model produced an R^2 value of .561 (adjusted $R^2 = .558$), $p < .001$. To investigate the predictive power of each variable, further multiple regression analyses were performed. In each case, all of the variables except the target variable were entered in the first step, to observe the change in R^2 when the target variable was added. Figure 1 shows the unique effect of each variable on naming RTs. The results

Table 1 Descriptive statistics of all predictors and naming RTs

Variable	Mean	Minimum	Maximum	Standard Deviation
z-scored RT	0.07	–0.83	1.83	0.52
Stop	0.25	0	1	0.43
Affricate	0.27	0	1	0.44
Fricative	0.19	0	1	0.39
Nasal	0.06	0	1	0.23
Liquid	0.10	0	1	0.30
Aspirated	0.41	0	1	0.49
Voiced	0.15	0	1	0.36
Bilabial	0.12	0	1	0.33
Labiodental	0.03	0	1	0.18
Alveolar	0.19	0	1	0.39
Palato-Alveolar	0.16	0	1	0.37
Alveo-Palatal	0.18	0	1	0.38
Velar	0.12	0	1	0.32
Freq	2.04	0.30	4.78	1.03
NS	13.17	4	30	4.07
CON	0.55	0	1	0.32
REG	0.46	0	1	0.50
UNP	0.11	0	1	0.31
PhonN	7.29	1	20	4.02
SemN	85.22	1	226	66.35
HD	1.06	0.30	1.85	0.31
SAR	0.40	0.28	0.72	0.08
FAM	4.09	1.25	6.75	1.11
NC1	0.48	0	2.08	0.45
NC2	0.43	0	2.10	0.45

Freq, character frequency; NS, number of strokes; CON, consistency; REG, regular; UNP, unpronounceable; PhonN, phonetic combinability; SemN, semantic combinability; HD, homophone density; SAR, semantic ambiguity rating; FAM, familiarity; NC1, number of disyllabic compound words formed by a character in the first position of the word; NC2, number of disyllabic compound words formed by a character in the second position of the word.

showed FAM to be the strongest predictor ($\Delta R^2 = 6.2$ %, $f^2 = 0.07$) when all other variables were entered into the model. CON was found to be the second strongest predictor ($\Delta R^2 = 1.9$ %, $f^2 = 0.02$), and REG the third ($\Delta R^2 = 0.7$ %, $f^2 = 0.007$). Other variables, including Freq ($\Delta R^2 = 0.3$ %, $f^2 = 0.003$), SAR ($\Delta R^2 = 0.2$ %, $f^2 = 0.002$), UNP ($\Delta R^2 = 0.1$ %, $f^2 = 0.001$), and NC ($\Delta R^2 = 0.1$ %, $f^2 = 0.001$), accounted for small but reliable portions of the variance in z-scored RTs, whereas NS, PhonN, SemN, HD, and NC2 did not have significant effects on predicting z-scored RTs.

Exploring an interaction term

To demonstrate the potential use of the present large-scale naming database, an additional multiple regression analysis

Table 2 Correlations between predictors (except initial phonemes) and naming RTs used in the regression model

		Freq	NS	CON	REG	UNP	PhonN	SemN	HD	SAR	FAM	NC1	NC2
IV	Freq	1											
	NS	−.141[**]	1										
	CON	−.038[*]	.110[**]	1									
	REG	−.127[**]	.009	.451[**]	1								
	UNP	.049[**]	.125[**]	.055[**]	−.321[**]	1							
	PhonN	.029	−.136[**]	−.381[**]	−.088[**]	−.100[**]	1						
	SemN	.014	−.181[**]	.131[**]	.087[**]	−.030	−.099[**]	1					
	HD	−.065[**]	.047[**]	.130[**]	.149[**]	.029	.015	.038[*]	1				
	SAR	.736[**]	−.137[**]	−.066[**]	−.154[**]	.043[*]	.022	.033	−.081[**]	1			
	FAM	.838[**]	−.175[**]	−.055[**]	−.140[**]	.027	.016	.016	−.133[**]	.704[**]	1		
	NC1	.790[**]	−.114[**]	−.088[**]	−.151[**]	.049[**]	.040[*]	.031	−.109[**]	.746[**]	.747[**]	1	
	NC2	.766[**]	−.088[**]	−.048[**]	−.123[**]	.042[*]	.015	−.040[*]	−.065[**]	.700[**]	.703[**]	.687[**]	1
DV	RT	−.618[**]	.089[**]	−.199[**]	−.071[**]	.032	.074[**]	−.068[*]	.059[**]	−.533[**]	−.669[**]	−.545[**]	−.521[**]

Freq, character frequency; NS, number of strokes; CON, consistency; REG, regular; UNP, unpronounceable; PhonN, phonetic combinability; SemN, semantic combinability; HD, homophone density; SAR, semantic ambiguity rating; FAM, familiarity; NC1, number of disyllabic compound words formed by a character in the first position of the word; NC2, number of disyllabic compound words formed by a character in the second position of the word. [*] Correlation is significant at the .05 level. [**] Correlation is significant at the .01 level.

was conducted to investigate whether the present database could demonstrate the interaction between consistency and character frequency that has typically been shown in previous studies with a factorial design (Lee et al., 2005; Yang et al., 2009). First the two variables, frequency and consistency, were centered, and then an interaction term was created by multiplying the values. Again, the initial phonemes were entered in Block 1. All other variables were entered into the model in Block 2, except the interaction term, which was included last. As can be seen in Table 4, the interaction term was significant (beta = .094, $p < .001$), showing that the consistency effect was influenced by character frequency.

According to Aiken and West (1991), a regression equation can be used to examine how the consistency effect of predicting RTs varies with frequency. All variables except frequency, consistency, and the interaction term were replaced by 0 in order to focus on the partial contributions from consistency and frequency and to illustrate how the predicted RTs varied as a function of consistency with the level of frequency. The equation could be formulated as follows:

$$Y = 1.322 - 0.259 \times (\text{consistency} - 0.55) - 0.058$$
$$\times (\text{frequency} - 2.04) + 0.152 \times \text{Interaction}.$$

We examined three different levels of frequency: high (one standard deviation above average frequency), medium (average frequency), and low (one standard deviation below average frequency).

As can be seen in Fig. 2, the regression slopes of the z-scored RTs for consistency in the low-, medium-, and high-frequency conditions all significantly differed from zero ($ps < .01$). However, the slope for low frequency (beta = −.416) was steeper than that for medium frequency (beta = −.259), and the slope for medium frequency was steeper than that for high frequency (beta = −.102), showing that the consistency effect was most prominent in the low-frequency condition.

Discussion

The data presented in this study include the lexical properties of 3,314 characters and their average responses, collected from 140 participants in a naming task. To the best of our knowledge, this is the first large-scale naming database for traditional Chinese characters.

Regression results showed that naming RTs were affected by frequency, consistency, and regularity. This is consistent with the data from behavioral (Lee et al., 2005), event-related potential (Hsu, Tsai, Lee, & Tzeng, 2009), functional magnetic resonance imaging (Lee et al., 2004), and computational modeling (Yang et al., 2009) studies. In particular, using dummy coding, the regression results showed that the naming RTs for characters with regular phonetic radicals and unpronounceable radicals were significantly faster than those for irregular phonetic radicals. The regression results also showed that initial phonemes and familiarity were significant predictors, whereas homophone density was not a significant predictor of naming RTs. Similar results had been reported in the regression data on a large-scale dataset of simplified Chinese characters by Liu et al. (2007).

Table 3 Regression coefficients of all predictors

Predictor	Naming			
	Beta	df	p Value	Effect Size (f^2)
Block 1				
Stop	.180		<.001***	
Affricate	.103		<.001***	
Fricative	.152		<.001***	
Nasal	–.118		.292	
Liquid	–.175		.213	
Aspirated	–.073		<.001***	
Voiced	.259		.130	
Bilabial	–.160		<.001***	
Labiodental	–.002		.892	
Alveolar	–.166		<.001***	
Palato-Alveolar	–.007		.740	
Alveolo-Palatal	.033		.127	
Vela	–.115		<.001***	
R^2 (%)	3.2 %	13		0.029
Block 2				
Freq	–.115		<.001***	
NS	.012		.349	
CON	–.173		<.001***	
REG	–.105		<.001***	
UNP	–.033		.011*	
PhonN	.011		.401	
SemN	–.019		.111	
HD	–.021		.104	
SAR	–.080		<.001***	
FAM	–.487		<.001***	
NC1	–.057		.007**	
NC2	–.026		.175	
R^2 (%)	56.1 %	25		1.28

The Beta coefficients presented are based on standardized scores. Freq, character frequency; NS, number of strokes; CON, consistency; REG, regular; UNP, unpronounceable; PhonN, phonetic combinability; SemN, semantic combinability; HD, homophone density; SAR, semantic ambiguity rating; FAM, familiarity; NC1, number of disyllabic compound words formed by a character in the first position of the word; NC2, number of disyllabic compound words formed by a character in the second position of the word. * Correlation is significant at the .05 level. ** Correlation is significant at the .01 level. *** Correlation is significant at the .001 level.

As was shown in Table 3, phonetic combinability did not exert a significant effect on naming RTs. Although Li et al. (2011) showed an inhibitory effect for naming characters with large phonetic combinability, the effect was only observed for low- and not for high-consistency characters. This might explain why the null effect of phonetic combinability was observed in the present regression analysis, because a wide range of consistency values were used. Further investigation could
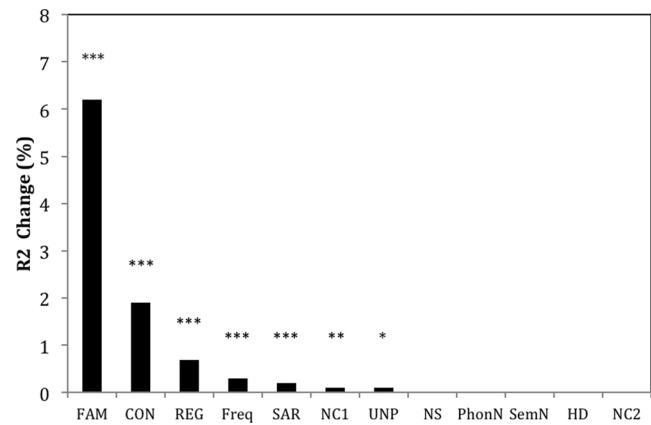
Fig. 1 Unique effect of each variable on naming reaction times. * Correlation is significant at the .05 level. ** Correlation is significant at the .01 level. *** Correlation is significant at the .001 level.

be conducted to explore the interaction between phonetic combinability and consistency. On the other hand, several studies have reported that greater semantic combinability could facilitate response latencies in either a semantic category task (Chen & Weekes, 2004) or a character decision task (Feldman & Siok, 1999). However, the regression results showed that semantic combinability was not a significant variable in predicting naming RTs. This discrepancy may have been due to differences in task demands: A semantic categorization task could benefit from access to knowledge about semantic radicals, but the naming task might not.

Another variable—number of strokes, measuring the visual complexity of characters—was not a significant predictor in the regression model. The finding of this null effect is consistent with the data reported by Su and Samuels (2010) in lexical decision, although they did observe an inhibitory effect for beginner readers. In addition, Leong, Cheng, and Mulcahy (1987) showed the effect only for low-frequency, but not for high-frequency characters in naming, and the effect was found in lexical decision when both high-frequency and low-

Table 4 Results from a three-block regression analysis for the exploration on the interaction term between frequency and consistency

Predictor	Naming			
	R^2 (%)	df	Effect Size (f^2)	p Value
Block 1: Initial phonemes	3.2	13	0.029	<.001***
Block 2: Lexical variables	56.1	25	1.28	<.001***
Block 3: Interaction	57.0	26	1.33	<.001**

The lexical variables included frequency, number of strokes, consistency, regular, unpronounceable, phonetic combinability, semantic combinability, homophone density, semantic ambiguity rating, familiarity, number of disyllabic compound words formed by a character in the first position of the word, and number of disyllabic compound words formed by a character in the second position of the word. ** Correlation is significant at the .01 level. *** Correlation is significant at the .001 level.
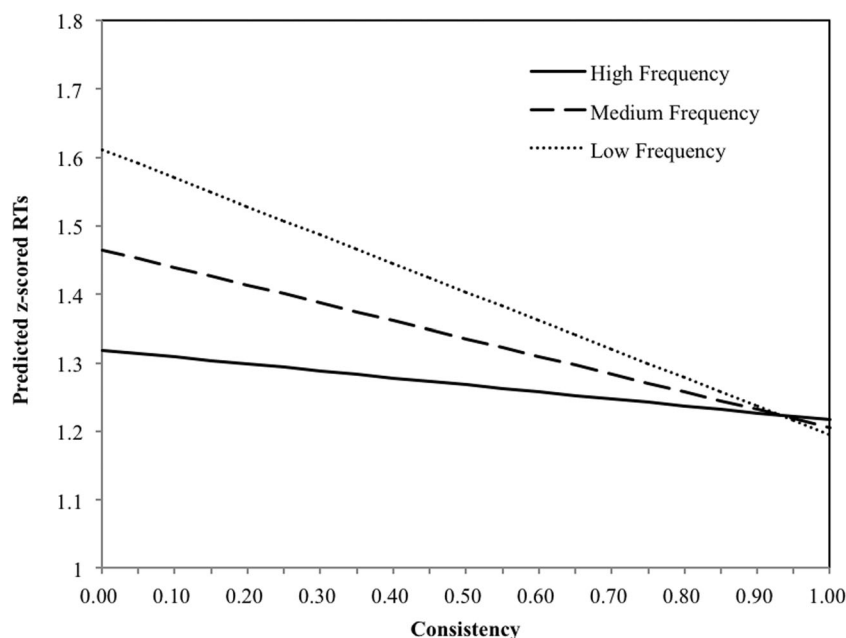
Fig. 2 Predicted *z*-scored reaction times (RTs) as a function of consistency in the high-frequency, medium-frequency, and low-frequency conditions.

frequency characters were mixed. These studies were all based on traditional Chinese characters. Studies using the simplified script—for instance, Liu et al. (2007)—have demonstrated a significant inhibitory effect of number of strokes in regression analyses of large-scale naming data. The discrepancy between previous studies and the present one seems to suggest that the effect of number of strokes is affected by objective frequency and subjective familiarity. It may be further complicated by different Chinese scripts; however, this would require further investigation.

In terms of the semantic variable, the semantic ambiguity rating accounted for a small but reliable portion of variance in the regression model. This result suggests that a character with multiple meanings is processed more efficiently, which is congruent with previous naming (Woollams, 2005) and lexical decision (Rodd, Gaskell, & Marslen-Wilson, 2002) studies.

In addition, the regression results showed that the index of the number of disyllabic compound words formed by a character in the first position of the word was significant, whereas the index of the number of disyllabic compound words formed by a character in the second position of the word only approached a significant level, suggesting that naming a character may automatically activate its corresponding disyllabic compound words, in particular when the character is embedded in the first position. This is consistent with the findings of studies using Chinese disyllabic compound words (Huang et al., 2006; Tsai, Lee, Lin, Tzeng, & Hung, 2006).

Finally, the effect of consistency varying with different levels of character frequency has been consistently reported by studies of Chinese character naming (Lee et al., 2005; Yang et al., 2009) that have used a factorial design. That is,

inconsistent characters take longer to name than consistent characters, especially for low-frequency characters. The regression analysis conducted here produced a similar interaction pattern (Fig. 2). Overall, the regression results showed the reliability of the present psycholinguistic database.

In summary, the primary aim of this study was to provide a large-scale psycholinguistic database for research into traditional Chinese language processing. The normative data were validated using multiple regression analyses to reproduce the effects of frequency, consistency, regularity, familiarity, and semantic ambiguity, and the interaction between frequency and consistency. The resulting database contains both lexical variables and naming RTs, which can be useful for psycholinguistic or educational research into Chinese character processing. The database can be accessed via an online inquiry system (http://ball.ling.sinica.edu.tw/namingdatabase/index.html). Three character-based queries are available: (1) to generate lists of items with specific lexical properties; (2) to submit lists of items in order to generate lexical properties; and (3) to submit a character in order to generate the lexical properties of its phonetic or semantic neighbors. The system also allows search results to be sorted by a specific lexical property and to be directly exported.

References

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. New York: Sage.

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General, 133,* 283–316. doi:10.1037/0096-3445.133.2.283

Balota, D. A., Pilotti, M., & Cortese, M. J. (2001). Subjective frequency estimates for 2,938 monosyllabic words. *Memory & Cognition, 29,* 639–647. doi:10.3758/BF03200465

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., & Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods, 39,* 445–459. doi:10.3758/BF03193014

Barca, L., Burani, C., & Arduino, L. S. (2002). Word naming times and psycholinguistic norms for italian nouns. *Behavior Reseach Methods, Instruments, & Computers, 34,* 424–434. doi:10.3758/BF03195471

Chen, M. J., & Weekes, B. S. (2004). Effects of semantic radicals on chinese character categorization and character decision. *Chinese Journal of Psychology, 46,* 181–196.

Coltheart, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology, 33A,* 497–505. doi:10.1080/14640748108400805

Fang, S.-P., Horng, R.-Y., & Tzeng, O. J. L. (1986). Consistency effects in the chinese character and pseudo-character naming tasks. In H. S. R. Kao & R. Hoosain (Eds.), *Linguistics, psychology, and the chinese language* (pp. 11–21). Hong Kong: Center of Asian Studies, University of Hong Kong.

Feldman, L. B., & Siok, W. W. T. (1999). Semantic radicals contribute to the visual identification of chinese characters. *Journal of Memory and Language, 40,* 559–576.

Ferrand, L., New, B., Brysbaert, M., Keuleers, E., Bonin, P., Méot, A., & Pallier, C. (2010). The French Lexicon Project: Lexical decision data for 38,840 French words and 38,840 pseudowords. *Behavior Research Methods, 42,* 488–496. doi:10.3758/BRM.42.2.488

Glushko, R. J. (1979). The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance, 5,* 674–691. doi:10.1037/0096-1523.5.4.674

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2009). *Multivariate data analysis* (7th ed.). New York, NY: Prentice Hall.

Hsu, C.-H., Lee, C.-Y., & Marantz, A. (2011). Effects of visual complexity and sublexical information in the occipitotemporal cortex in the reading of Chinese phonograms: A single-trial analysis with meg. *Brain and Language, 117,* 1–11.

Hsu, C.-H., Tsai, J.-L., Lee, C.-Y., & Tzeng, O. J. (2009). Orthographic combinability and phonological consistency effects in reading chinese phonograms: An event-related potential study. *Brain and Language, 108,* 56–66.

Huang, C.-R., & Chen, K.-J. (1998). Academia Sinica balanced corpus. Retrieved from www.sinica.edu.tw/SinicaCorpus/

Huang, C.-Y., Lee, C. Y., Huang, H.-W., & Chou, C.-J. (2011). Number of sense effects of chinese disyllabic compounds in the two hemispheres. *Brain and Language, 119,* 99–109.

Huang, H.-W., Lee, C. Y., Tsai, J. L., Lee, C. L., Hung, D. L., & Tzeng, O. J. (2006). Orthographic neighborhood effects in reading chinese two-character words. *NeuroReport, 17,* 1061–1065.

Hue, C.-W. (1992). Recognition processes in character naming. In H.-C. E. Chen & O. J. L. Tzeng (Eds.), *Language processing in Chinese (Advances in Psychology* (Vol. 90, pp. 93–107). Amsterdam, The Netherlands: North-Holland.

Jared, D. (1997). Spelling–sound consistency affects the naming of high-frequency words. *Journal of Memory and Language, 36,* 505–529.

Jared, D., McRae, K., & Seidenberg, M. S. (1990). The basis of consistency effects in word naming. *Journal of Memory and Language, 29,* 687–715.

Keuleers, E., Diependaele, K., & Brysbaert, M. (2010). Practice effects in large-scale visual word recognition studies: A lexical decision study on 14,000 Dutch mono- and disyllabic words and nonwords. *Frontiers in Psychology, 1*(174), 1–15. doi:10.3389/fpsyg.2010.00174

Keuleers, E., Lacey, P., Rastle, K., & Brysbaert, M. (2012). The British Lexicon Project: Lexical decision data for 28,730 monosyllabic and disyllabic English words. *Behavior Research Methods, 44,* 287–304. doi:10.3758/s13428-011-0118-4

Lee, C.-Y., Huang, H.-W., Kuo, W.-J., Tsai, J.-L., & Tzeng, J.-L. O. (2010). Cognitive and neural basis of the consistency and lexicality effects in reading Chinese. *Journal of Neurolinguistics, 23,* 10–27. doi:10.1016/j.jneuroling.2009.07.003

Lee, C.-Y., Tsai, J. L., Kuo, W. J., Yeh, T. C., Wu, Y. T., Ho, L. T., & Hsieh, J. C. (2004). Neuronal correlates of consistency and frequency effects on chinese character naming: An event-related fMRI study. *NeuroImage, 23,* 1235–1245.

Lee, C.-Y., Tsai, J.-L., Su, E. C.-I., Tzeng, O. J., & Hung, D. L. (2005). Consistency, regularity, and frequency effects in naming chinese characters. *Language and Linguistics, 6,* 75–107.

Leong, C. K., Cheng, P.-W., & Mulcahy, R. (1987). Automatic processing of morphemic orthography by mature readers. *Language and Speech, 30,* 181–196.

Li, Q.-L., Bi, H.-Y., Wei, T.-Q., & Chen, B.-G. (2011). Orthographic neighborhood size effect in Chinese character naming: Orthographic and phonological activations. *Acta Psychologica, 136,* 35–41.

Liu, Y., Shu, H., & Li, P. (2007). Word naming and psycholinguistic norms: Chinese. *Behavior Research Methods, 39,* 192–198. doi:10.3758/BF03193147

McBride-Chang, C., Chow, B. W. Y., Zhong, Y., Burgess, S., & Hayward, W. G. (2005). Chinese character acquisition and visual skills in two Chinese scripts. *Reading and Writing, 18,* 99–128.

Ministry of Education, Taiwan. (2012). *Table of Mandarin polyphon.* Taiwan: Mandarin Promotion Council, Ministry of Education, Taiwan. Retrieved from www.edu.tw/pages/detail.aspx?Node=3692&Page=16373&WID=c5ad5187-55ef-4811-8219-e946fe04f725

Murray, W. S., & Forster, K. I. (2004). Serial mechanisms in lexical access: The rank hypothesis. *Psychological Review, 111,* 721–756. doi:10.1037/0033-295X.111.3.721

New, B., Ferrand, L., Pallier, C., & Brysbaert, M. (2006). Reexamining the word length effect in visual word recognition: New evidence from the english lexicon project. *Psychonomic Bulletin & Review, 13,* 45–52. doi:10.3758/BF03193811

Perfetti, C. A., & Tan, L. H. (1998). The time course of graphic, phonological, and semantic activation in Chinese character identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 101–118. doi:10.1037/0278-7393.24.1.101

Perry, C., Ziegler, J. C., & Zorzi, M. (2007). Nested incremental modeling in the development of computational theories: The CDP+ model of reading aloud. *Psychological Review, 114,* 273–315. doi:10.1037/0033-295X.114.2.273

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103,* 56–115. doi:10.1037/0033-295X.103.1.56

Rodd, J., Gaskell, G., & Marslen-Wilson, W. (2002). Making sense of semantic ambiguity: Semantic competition in lexical access. *Journal of Memory & Language, 46,* 245–266. doi:10.1006/jmla.2001.2810

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96,* 523–568. doi:10.1037/0033-295X.96.4.523

Su, Y.-F., & Samuels, S. J. (2010). Developmental changes in character-complexity and word-length effects when reading Chinese script. *Reading and Writing, 23,* 1085–1108.

Sze, W. P., Rickard Liow, S. J., & Yap, M. J. (2013). The Chinese Lexicon Project: A repository of lexical decision behavioral responses for 2, 500 Chinese characters. *Behavior Research Methods, 46,* 263–273. doi:10.3758/s13428-013-0355-9

Tan, L. H., Hoosain, R., & Siok, W. W. T. (1996). Activation of phonological codes before access to character meaning in written Chinese. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 865–882. doi:10.1037/0278-7393.22.4.865

Taraban, R., & McClelland, J. L. (1987). Conspiracy effects in word pronunciation. *Journal of Memory and Language, 26,* 608–631.

Tsai, J.-L., Lee, C. Y., Lin, Y. C., Tzeng, O. J., & Hung, D. L. (2006). Neighborhood size effects of chinese words in lexical decision and reading. *Language and Linguistics, 7,* 659–675.

Woollams, A. M. (2005). Imageability and ambiguity effects in speeded naming: Convergence and divergence. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 878–890. doi:10.1037/0278-7393.31.5.878

Yang, J., McCandliss, B. D., Shu, H., & Zevin, J. D. (2009). Simulating language-specific and language-general effects in a statistical learning model of chinese reading. *Journal of Memory and Language, 61,* 238–257.

Yap, M. J., Rickard Liow, S. J., Jalil, S. B., & Faizal, S. S. B. (2010). The Malay Lexicon Project: A database of lexical statistics for 9,592 words. *Behavior Research Methods, 42,* 992–1003. doi:10.3758/BRM.42.4.992

Zhou, Y. G. (1978). To what degree are the "phonetics" of present day Chinese characters still phonetic? *Zhongguo Yuwen, 146,* 172–177.