Behavior Research Methods (2024) 56:4732–4757
https://doi.org/10.3758/s13428-023-02214-1

A large-scale database of Chinese characters and words collected from elementary school textbooks

Man Zhang[1] · Zeping Liu[2] · Mona Roxana Botezatu[3] · Qinpu Dang[1] · Qiming Yuan[1] · Jinzhuo Han[4] · Li Liu[1] · Taomei Guo[1]

Accepted: 30 July 2023 / Published online: 24 August 2023

Abstract

Lexical databases are essential tools for studies on language processing and acquisition. Most previous Chinese lexical databases have focused on materials for adults, yet little is known about reading materials for children and how lexical properties from these materials affect children's reading comprehension. In the present study, we provided the first large database of 2999 Chinese characters and 2182 words collected from the official textbooks recently issued by the Ministry of Education (MOE) of the People's Republic of China for most elementary schools in Mainland China, as well as norms from both school-aged children and adults. The database incorporates key orthographic, phonological, and semantic factors from these lexical units. A word-naming task was used to investigate the effects of these factors in character and word processing in both adults and children. The results suggest that: (1) as the grade level increases, visual complexity of those characters and words increases whereas semantic richness and frequency decreases; (2) the effects of lexical predictors on processing both characters and words vary across children and adults; (3) the effect of age of acquisition shows different patterns on character and word-naming performance. The database is available on Open Science Framework (OSF) (https://osf.io/ynk8c/?view_only=5186bd68549340bd923e9b6531d2c820) for future studies on Chinese language development.

Keywords Chinese · Lexical database · Elementary school textbooks · Word-naming task

Introduction

Lexical databases not only provide a rich array of lexical properties from orthographic, phonological, and semantic perspectives for researchers to access and retrieve, but also facilitate research in lexical processing by providing high-quality and reliable normative data. Despite the fact that several databases on Chinese language, a logographic writing system in which a written character usually represents a word or a semantic unit (De Francis, 1989; Yang et al., 2009), have been constructed over the past two decades (Liu et al., 2007; Cai & Brysbaert, 2010; Sze et al., 2014; Chang et al., 2016; Tse et al., 2017; Tsang et al., 2018; Sun et al., 2018; Wang et al., 2020), resources based on reading materials for school-age children in Mainland China are surprisingly limited (Shu et al., 2003; Xing et al., 2004; Cai et al., 2021; Li et al., 2022). Most recently, Li et al., (2022) has developed a lexical database that incorporates Chinese characters and words sampled from elementary school children's curricular and extracurricular books and provides frequency statistics. Nonetheless, it remains to be explored how various lexical properties of these characters and words change with grade levels, and how these factors affect lexical processing among readers of different age.

To fill these gaps, in the present study we aim to establish a database for Chinese characters and words collected from the official textbooks recently issued by the Ministry of Education of the People's Republic of China (MOE) for most elementary schools in Mainland

✉ Taomei Guo
guotm@bnu.edu.cn

[1] State Key Laboratory of Cognitive Neuroscience and Learning & IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing 100875, China

[2] Department of East Asian Languages and Cultures, Indiana University, Bloomington, IN, USA

[3] Department of Speech, Language and Hearing Sciences, University of Missouri, Columbia, MO, USA

[4] Chinese Language and Culture College, Beijing Normal University, Beijing 100875, China

China, and further identify the trajectory of lexical factors and its effects on Chinese literacy acquisition. In the following sections of the introduction, we first present an overall picture of Chinese lexical units followed by a summary of existing databases. We then focus on the source of our database – the current textbooks used in most elementary schools in Mainland China. Finally, we present a detailed description of the lexical variables involved in our database.

Chinese lexical units

Most words in Mandarin Chinese are represented in disyllabic forms, i.e., with two characters (He & Li, 1987). Chinese characters are fundamental building blocks of a Chinese lexicon. Taking a close look at the internal structure of each Chinese character, one can find that it is composed of various radicals, which are also formed by different strokes arranged in a stipulated order (DeFrancis, 1989). For example, the character "江", meaning *river*, consists of two radicals: the left radical "氵" and the right radical "工". "工" is also composed of two basic strokes: the horizontal "一" and the vertical "丨". Therefore, the writing system of Chinese is organized in a hierarchical way: from words to characters, to radicals, to strokes (McBride, 2016).

It has been shown in previous studies that learning Chinese characters is affected by the combination and distribution of different sub-lexical levels, such as radicals or strokes. For example, Tong and McBride (2014) found that children had already developed the positional awareness (i.e., the ability to know the positional constraints of radicals or stroke patterns) as early as in kindergarten and were able to name pseudo-characters after simple instruction. Moreover, the diverse and flexible combination of radicals in Chinese is partly due to the fact that approximately 80% of characters are compound characters, which are composed of two or more radicals, according to DeFrancis (1989) and Zhu (1988). Most compound characters include two critical radicals: a semantic radical that provides lexical-semantic cues for a character, and a phonetic radical that shares pronunciation information with a character (Li et al., 2022; Sun et al., 2018; Wang et al., 2020). Take "洋" (yáng, meaning *ocean*) as an example. The semantic radical, '氵' indicates that this character is related to *water*, while the phonetic radical, '羊' on the right side, shares the same pronunciation with '洋'. These compound characters are termed phonograms in the literature. Nevertheless, Fan et al. (1984) suggested that only a limited number (about 26%) of phonograms could bear phonological similarity with their phonetic radicals. Furthermore, semantic radicals have varied transparency (referring to the extent to which the lexical meaning of the character can be derived from its semantic radical) and combinability (also known as neighborhood density, referring to the number of compound characters that share the same semantic radicals; Chen et al., 2006; Feldman & Siok, 1999; McBride, 2016; Shu et al., 2003; Tzeng et al., 1995). Knowing the regularity or pattern embedded in phonograms facilitates lexical recognition and literacy acquisition (Ho et al., 2003; Packard et al., 2006).

Existing Chinese lexical databases

As more researchers have begun to realize the importance of lexical databases for psycholinguistic research, several large-scale Chinese lexical databases have been published. Table 1 summarizes 12 Chinese lexical databases in the past two decades. The materials of four databases on written language were extracted from various corpora published previously and three databases were based on sampled characters and/or words from Cai & Brysbaert (2010), a frequency database for spoken language collected from film and television subtitles. Although they differ in sources and materials, most of them included normative data from adults and most relevant factors that affect Chinese character and word processing. To achieve this, these databases included a wide range of lexical variables such as phonology, orthography, semantics, and frequency to examine their effects in lexical processing tasks. Note that some databases not only have expanded the number of Chinese characters and words but have also delved into more precise classifications while taking into account the relationships among characters. Chang et al. (2016), for example, examined the consistency effect (i.e., the number of characters sharing the same phonetic radicals and have the same sound) on naming RTs and found that this effect was modulated by character frequency. Furthermore, Tsang et al. (2018) and Sun et al. (2018) extended their investigations by including three-character words or even four-character words, which basically encompass all constructive units in Chinese, offering the possibility to explore the relationship between words and characters.

Among these studies, lexical decision and word-naming tasks are most frequently used to test the effects of various lexical characteristics in word identification (Katz, et al., 2012). Most recently, Wang et al. (2020) have adopted dictation task to study the contribution of 14 lexical variables on 1600 characters to writing measurements, including writing latency, duration, and accuracy, which provides a valuable resource for studies on Chinese handwriting. In the present study, the normative data were collected through the word-naming task as Liu et al. (2007) and Chang et al. (2016).

It should be noted that most existing norms of the above-mentioned Chinese databases are exclusively obtained from materials for adults, so they may not be suitable for studies on school-aged children. On the other hand, children-centered databases have been well established in most alphabetical languages, such as the *Children's Printed Word Database*

Table 1 Published Chinese lexical databases over the last 20 years

Database	Purpose	Material	Source	Lexical Variables	Task
Shu et al., (2003)	To understand children's acquisition of characters	2570 simplified characters	Elementary school textbooks (1996)	Visual complexity, phonetic regularity and consistency, phonetic and semantic families, etc	No
Liu et al., (2007)	To examine potential variables that contribute to word-naming performance	2423 simplified characters	Language Corpus System of Modern Chinese Studies (LCSMCS; Sun et al., 1997)	Frequency, homophone density, AoA, familiarity, concreteness, imageability, regularity, etc	Word naming
Cai & Brysbaert (2010)	To build a database of word and character frequencies	5936 characters and 99,121 words, simplified	A corpus of film and television subtitles	CHRCount, CHR / million, logCHR and CHR-CD, CHR-CD%, logCHR-CD	Lexical decision
Sze et al. (2014)	To develop a Chinese Lexicon Project	2500 simplified characters	Lu (1989, 1992) teaching corpus	Focused on Frequency and conducted 3 virtual simulations including number of stokes, meaning, and AoA	Lexical decision
Chang et al. (2016)	To provide a large-scale set of psycholinguistic norms	3314 traditional characters	Academia Sinica Balanced Corpus (ASBC; Huang & Chen, 1998)	Frequency, regularity, semantic ambiguity rating, phonetic combinability, etc	Word naming
Tse et al., (2017)	To provide norms for Chinese compound words	25,000+ traditional words	A pocket dictionary (Que, 2008) and the suggested word pool in Microsoft Word database	Number of strokes and frequency from various sources	Lexical decision
Tsang et al., (2018)	To develop a mega-study for simplified Chinese words	1020 one-character, 10,022 two-character, 949 three-character, 587 four-character, simplified	SUBTLEX-CH (Cai & Brysbaert, 2010)	Length, number of stokes and frequency, etc	Lexical decision
Sun et al., (2018)	To develop a Chinese Lexical Database for single and multi-character words	3913 one-character words, 34,233 two-character words, 7143 three-character words, 3355 four-character words, simplified	SUBTLEX-CH (Cai & Brysbaert, 2010) and Leiden Weibo Corpus (Van Esch, 2012)	Frequency, complexity, neighborhood density, orthography-phonology consistency, information-theoretic measures, etc	Lexical decision, word-naming
Wang et al., (2020)	To provide a psycholinguistic database of Chinese character handwriting	1600 simplified characters	SUBTLEX-CH (Cai & Brysbaert, 2010)	Frequency, meaning, AoA, phonogram, regularity, structure, etc	Spelling-to-dictation task
Xu et al., (2020)	To provide Age of Acquisition (AoA) norms for Chinese words	19,716 simplified Chinese words	Language Situation in China, published by the National Language Commission (NLC), Ministry of Education, PRC (2011–2018); MELD-SCH (Tsang et al., 2018)	Subjective AoA rated by native speakers of Mandarin Chinese	To recall the age range in which a word is learned
Cai et al., (2021)	To provide two AoA norms based on the textbooks	3300+ simplified Chinese characters	Chinese textbooks, published by the People's Education Press in 2001 and 2011	Objective AoA based on Chinese school textbooks	No

Table 1 (continued)

Database	Purpose	Material	Source	Lexical Variables	Task
Li et al., (2022)	To provide grade-level frequency and contextual diversity statistics for Chinese characters and words	6746 simplified characters and 153,079 words	Curricular books published by the People's Education Press and popular extracurricular books (Primary and Middle School)	Frequency, contextual diversity, word length and part-of-speech category	Word naming

Word-naming task: Participants read aloud all characters or words as quickly and accurately as possible. Lexical decision task: Participants judge whether the stimulus is a real Chinese character/word or not. Spelling-to-dictation task: Participants write down the character according to spoken phrase.

(Masterson et al., 2010; Stuart et al., 2003) for British English; NOVLEX (Lambert & Chesnet, 2001) and MANULEX (Lété et al., 2004) databases for French; LEXIN(Corral et al., 2009) and ONESC (Martín & Pérez, 2008) databases for Spanish; Lessico Elementare (Marconi et al., 1993) for Italian; childLex (Schroeder et al., 2015) for German; and ESCOLEX (Soares et al., 2014) for Portuguese.

Shu et al.'s (2003) was the first study to investigate various properties of Chinese characters and their distributions at each grade level based on elementary school textbooks (1996) issued by the Ministry of Education used in Beijing and other regions at that time. Specifically, they focused on the development of certain features, e.g., visual complexity and the proportion of regular character (i.e., characters containing phonetic radicals that can provide salient clues to the pronunciation of the character), across different grades. This database indeed has been a useful tool for studies on Chinese language acquisition in children (e.g., Chen et al., 2009a, b; Li et al., 2012, 2020; Song et al., 2015; Tong et al., 2009). Xing et al. (2004) also constructed a corpus of Chinese characters from the textbooks issued more than two decades ago (Beijing Academy of Educational Sciences, 1998). They examined the role of regularity, consistency, and frequency in Chinese character processing, and confirmed the effects of these lexical variables in the acquisition of Chinese characters. Other databases established on textbooks can be found in Cai et al. (2021), which include new age of acquisition (AoA) norms for 3300+ simplified Chinese characters based on Chinese textbooks of the 2001 and the 2011 editions, both of which were published by the People's Education Press and were once widely used in elementary schools in mainland China. Moreover, they found that the objective AoA norms from these textbooks can well explain accuracy and reaction times in lexical access tasks used in previous databases. Given the fact that textbooks for elementary schools in mainland China have been undergoing changes in recent decades, it is necessary to develop new databases based on newly issued textbooks. Therefore, in the present study, we aim to establish a more up-to-date Chinese lexical database from children's learning materials. Such an attempt can help to uncover how the properties of Chinese lexical units evolve across different developmental stages and provide an effective tool for studies on Chinese language learning.

The instructional materials

In mainland China, the instructional materials used in elementary schools (i.e., Chinese language textbooks) are the primary source for children to receive formal language education. The series of Chinese language textbooks have undergone more than 70 years of revisions, from the initial *Mandarin Chinese Textbook*, which was revised and adapted by the North China Textbook Editorial Committee in 1949, to the most recent *Chinese Compulsory Education Textbook*, which was validated by the MOE in 2017 (Li, 2018; Wang & Chen, 2019). Textbooks have varied greatly from version to version over the past several decades in terms of (1) requirements and (2) selections for texts and words. Issued in the fall of 2017, the new edition of the textbooks, which places more emphasis on the scientific arrangement of the content, was introduced for use nationwide. These textbooks are the first version that begins with learning Chinese characters rather than Pinyin (the official romanization scheme for Mandarin used in mainland China and some other Chinese-speaking regions) in the first volume, demonstrating that this set of textbooks attaches great importance to character learning. Moreover, there are a total of 12 volumes of the textbooks, ranging from the first to the sixth grade, with two volumes for each grade level. There are nearly 20 pieces of texts distributed across eight units for each volume. The additional section *Learning Activity* summarizes and extends basic information for each unit to help students enhance learning.

According to Wen (2016, 2017), the editor-in-chief of this edition, the textbooks were compiled based on the principle of "a separation of recognition and writing, more recognition and less writing". This principle foregrounds the distinction of ‘learning to read’ and ‘learning to write’ in that they set different goals for children; ‘learn to read’ is less demanding than ‘learn to write’ because the latter one requires children to write Chinese characters in stipulated sequences, which calls for a higher level of memory and motor schema (Zhang et al., 2021). Furthermore, the textbooks emphasize the need to learn Chinese characters through a combination of Pinyin, structure, and components. It may be worth noting that Chinese orthography places different demands on school-age children at different stages to acquire the knowledge of orthography, phonology, and the meaning of characters. For example, children in early grades are provided characters of simple structure and high frequency. As their orthographic awareness becomes mature in later grades, they begin to consolidate those already-learned characters and learn new characters and words with relatively low frequency. Accordingly, the 300 fundamental characters which are necessary for reading are provided in the first volume, reflecting the influence of phonological mediation, orthographic rules, and frequency effects on lexical access from a psycholinguistic perspective. Li (2018) also pointed out that Chinese characters that children need to write in the early years are generally simple characters with strong productivity so as to form other compound characters or words. In this regard, the learning materials can reasonably corroborate the evidence of children’s learning trajectory during their literacy development. Therefore, the present study aimed to examine the objective linguistic attributes of Chinese lexical units based on this set of textbooks. This attempt will provide more evidence to support the development and improvement of learning materials.

The current database

The current database contains a total of 2999 unique Chinese characters and 2182 words from textbooks for Grades 1 to 6. We incorporated multilevel lexical features regarding orthography, phonology, meaning, and frequency. The norming data collected from the word-naming task (primarily reaction times and accuracy) among school-aged children and adults were also included in order to examine the influential factors of reading in Chinese. Compared with previous databases, this study extends the current body of database in at least three aspects.

First, the current database systematically investigated the new version of instructional materials for elementary school children. Notably, in order to make suggestions for teaching Chinese language, we tested whether the distribution of characters and words across all grades was truly based on changes in some important lexical attributes. Also, we can use the data from the word-naming task to examine the contribution of these lexical variables to word recognition and production.

Second, the database is important for studies on the effect of age of acquisition (AoA). Numerous studies have demonstrated that words acquired at an earlier age have a significant advantage in word recognition and production (Bonin et al., 2001; Carroll & White, 1973; Gilhooly & Logie, 1980; Johnston & Barry, 2006; Pérez, 2007). However, most previous studies on AoA have relied on subjective measures, which require adult subjects to determine the age when they were able to speak or read a certain word. Such subjective recalling is easily confounded with other factors such as word frequency, familiarity, and/or complexity (Morrison et al., 1997; Morrison & Ellis, 2000) and therefore raises concerns about its validity (Stadthagen-Gonzalez & Davis, 2006). Due to the fact that most Chinese characters and words are acquired in school at early ages, a more objective way to define the AoA is the grade in which a character or a word first appears in textbooks (Cai et al., 2021; Shu et al., 2003; Wang et al., 2020). It should be noted that the textbook-based estimation of AoA centers on written characters and words, which differ from spoken forms. As Cai et al., (2021) indicated, children may have been exposed to spoken forms before school age. Moreover, the acquisition of spoken forms does not entail the acquisition of written forms of characters and words. Therefore, it bears reiterating that the AoA collected in the current database (represented by Volume) reflects the age of learning written forms of Chinese characters and words.

Third, patterns generalized from the current database can help children learn Chinese characters more efficiently. Previous studies have suggested that children are sensitive to statistical structures in a given language environment and are capable of extracting embedded abstract regularities to acquire languages (known as statistical learning, see Perruchet & Pacton, 2006; Romberg & Saffran, 2010; Saffran et al., 1996). Such ability has also been demonstrated to have associations with children's literacy outcomes (Arciuli & Simpson, 2012; Spencer et al., 2015). In character learning, statistical learning is manifest in many aspects, such as knowing the pronunciations of phonograms by consulting their phonetic radicals and developing awareness of visual orthographic regularities. The current database has provided information regarding to what extent the phonetic radical can inform the sound of character by categorizing characters into different types.

Lexical variables

Previous studies have shown that visual complexity, the number of different pronunciations and meanings, and character frequency play important roles in Chinese word recognition and production (Chang et al., 2016; Hsu et al., 2011; Liu et al., 2007; Peng et al., 2003; Peng & Wang, 1997; Tan & Perfetti, 1997; Wang et al., 2020). The multilevel lexical variables presented in the current database focus on four main aspects of Chinese characters and words: orthography, phonology, meaning, and frequency.

Strokes and radicals As the word length effect found in alphabetic languages, the stroke number effect detected in considerable studies on Chinese word recognition also suggests that response latency increases with the number of constituent strokes (Chen et al., 1996; Fang, 1994; Leong et al., 1987; Su & Samuels, 2010; Tan & Peng, 1990). The number of strokes in the current database was extracted from *Xinhua Dictionary* (新华字典, Linguistics Institute of Chinese Academy of Social Sciences, 2020), and the information about character radicals was obtained from *Dictionary of Chinese character properties* (汉字属性字典, Fu, 1989). The number of radicals is also a common proxy to measure visual complexity of Chinese characters, as indicated in previous studies (e.g., Liu et al., 2007; Wang & Dong, 2013; Xing et al., 2004). To ensure the number of radicals could also reflect the visual complexity of characters, we dissected the characters into basic radicals and then calculated the total number of radicals (Wang et al., 2020). Take '别' (bié, *other*) as an example: The basic radicals are '口', '力' and ' 刂' rather than '另'and ' 刂', so the number of radicals is 3.

Meaning and pronunciation Some characters such as '水' (shuǐ, *water*) and '人' (rén, *human beings*) convey specific meanings. Alternatively, many characters have multiple meanings and thus are semantically vague and highly context-dependent (Perfetti & Tan, 1998; Tan et al., 1995). For example, the character '打' (dǎ) means '*to get or obtain*' when it combines with the character '水' as '打水'(dǎshuǐ); but when it precedes the character '球' (qiú, *ball*), it means '*to play*'. Many studies have demonstrated that semantic ambiguity has an effect on word recognition (Borowsky & Masson, 1996; Kellas et al., 1988; Rodd et al., 2002).

In addition, some Chinese characters may have different pronunciations with different combinations of other characters (known as heteronyms). We also incorporated this attribute into the database to see how the variability of pronunciations affects lexical access. For example, the character '好' has two pronunciations that differ in the lexical tone: *hǎo* and *hào*, representing different syntactic categories. When it is pronounced as *hǎo*, it is an adjective that has six different meanings including 'good', 'friendly or harmony', and so on. When it is *hào*, it is a verb meaning 'to love or like'. We retrieved the meaning and pronunciation information for all characters from *Xinhua Dictionary* (新华字典, Linguistics Institute of Chinese Academy of Social Sciences, 2020). For words, we resorted to *Modern Chinese Dictionary* (现代汉语词典, Chinese Academy of Social Sciences, 2012), *The Great Chinese Word Dictionary* (汉语大词典; Luo, 1986) and *Xinhua Idiom Dictionary* (新华成语词典, Lexicographical Center of Commercial Press, 2002). However, 87 words were not found in these dictionaries and were treated as missing values in our analysis. Note that only literal meanings of characters and words were used in the current study. We did not consider the sense of characters and words that goes beyond the dictionary definitions, such as figurative, sarcastic, or ironic usages because children mainly learn the most basic meanings of these characters in elementary school stages.

Structure The relative positions of constituent radicals within characters also affect character recognition (Feldman & Siok, 1999; Li et al., 2012, 2000; Yang et al., 2019). In the current database, we identified six primary structures to capture the relationship of main radicals within a particular character: single, right-left, top-down, full-surrounded, semi-surrounded, and others (including two special structures, one of which is arranging three identical components, like '品' or '晶'; the other one is 'embedded structure', like '爽' or '巫'). Considering the vast majority (80%) of characters in our database are left-right (1744) and top-down (748) structure, the character structure was reclassified into three categories (i.e., left-right, top-down, and other) in the analysis below.

Type Traditionally, all Chinese characters can be classified into six basic types based on the manners in which they are formed and derived according to Xu Shen's Shuowen, also known as *liùshū* ('six writings'). Characters within the type of Xiàngxíng (‘pictographs’) are physically similar to the entities or objects that they represent, like ‘日’ (roughly resembling the shape of *the sun*). Zhǐshì is the other type that denotes ideas by indicating things in a metaphorical way; for example, ‘上’ with the focus on the upper part, is later used to indicate a higher level or upward. Apparently, they have a high degree of iconicity as they are either depicting or indicating entities or parts. Another two categories, Xíngshéng (phono-semantic compounds, also referred to as ‘picto-phonetic characters’ or ‘phonograms’) and Huìyì (associative compounds or compound ideographs), involve the combination of character radicals. The former one is composed of semantic and phonetic radicals (Li, 1993; Yin & Rohsenow, 1994), whereas the latter one is composed of pictographic or ideographic characters to convey the meaning. The last two types are actually associated with character etymology, Zhuǎnzhù (derivative cognates) and Jiǎjiè (rebus or phonetic loan characters; Myers, 2019), so they were excluded from the current database. The criterion for type of characters in the current database is based on online dictionaries ChaZiWang (http://www.chaziwang.com/) and Hanzi Quanxi Ziyuan Yingyong Xitong (https://qxk.bnu.edu.cn/). Considering the vast majority of compound characters existing in the system, we only distinguished whether a character belongs to the type of phonogram in our statistical analysis.

Regularity The ability of mapping letters to sounds is essential in reading acquisition (Brady & Shankweiler, 1991; Byrne, 1992; Goswami & Bryant, 1990). Regular words, which are in accordance with the GPC rules in alphabetic languages, have been demonstrated to be processed faster and more accurately compared to irregular words (Baron & Strawson, 1976; Parkin, 1982; Seidenberg et al., 1984). The correspondence between the written form and the sound in Chinese mostly manifests in phonograms, as we outlined earlier. Thus, we made a further distinction regarding the degree to which phonetic radicals inform the pronunciation of a character in order to obtain the patterns of regularity for these characters. Table 2 displays six types of phonological relations between compound characters and their phonetic radicals. The classification was adapted from Zhou and Marslen-Wilson (1999), with an additional type ‘unpronounceable’ added. For characters or phonetic radicals with multiple pronunciations, all pronunciations were considered. For example, the radical ‘隹’ has three pronunciations: zhuī, cuī and wéi. The character ‘准’ (zhǔn) shares the initial with one of the pronunciations of ‘隹’, thus ‘准’ is categorized as Alliteration. In addition, some radicals are presented in a different way when they serve as constituent radicals and independent characters; for instance, the upper part of the character ‘党’ (dǎng, *party*) is also a character ‘尚’(shàng, *still*), but it has been distorted and less likely to be recognized. Therefore, we treated those characters with distorted constituent radicals as “unpronounceable”. In our analysis below, three major types were investigated: regular, semiregular (including semi-regular, rhyming, and alliteration), and irregular (including both irregular and unpronounceable).

logCHR-CD and logW-CD Frequency effect, which has been observed consistently in a wide range of tasks across all languages, is a reliable and fundamental predictor in lexical access. In Chinese, Sze et al. (2014) summarized seven major Chinese character frequency norms established in

Table 2 Phonological relations between the compound character and its phonetic radical

Type	Description	Example
Regular	The character and its phonetic radical share the same pronunciation	青 – 清 qīng – qīng blue – clear
Semi-regular	The character and its phonetic radical share the same syllable but not lexical tone	票 – 飘 piào – piāo ticket - blow
Rhyming	The character and its phonetic radical share the same final	干 – 汗 gān – hàn dry – sweat
Alliteration	The character and its phonetic radical share the same initial (consonant)	某 – 煤 mǒu – méi certain – coal
Irregular	The phonetic radical cannot provide a cue regarding the character’s pronunciation	乃 – 仍 nǎi – réng be – still
Unpronounceable	The phonetic radical is unpronounceable and not an independent character	㐬 – 流 / – liú / – flow

the literature and suggested that the character frequency based on contextual diversity (CHR-CD, referring to the number of films in which a character occurs) from Cai & Brysbaert (2010) could account for the most variance (nearly 31%) in response times for the lexical decision in Chinese. The effect of character and word contextual diversity (CD) was also found in lexical processing in fourth-grade children (Huang, et al., 2021). Tse et al. (2017) compared six word-frequency measures in lexical decision performances and also observed the same patterns. Several subsequent studies used the subtitle frequencies from Cai and Brysbaert (2010) in their analyses (Sun et al., 2018; Tsang et al., 2018; Tse et al., 2017; Wang et al., 2020). In the current database, we included the logarithmic CHR-CD for all characters (represented as logCHR-CD) and logW-CD (Cai & Brysbaert, 2010) for all words to represent lexical frequency. Moreover, we included the contextual diversity (log-transformed) from CCLOWW in Li et al., (2022) for all characters and words to represent children's frequencies, which contrasts with adult frequencies from Cai and Brysbaert (2010). This variable was excluded from analysis (only reported in the descriptive statistics) due to the high collinearity with other frequency measures.

Count_Sum In addition to the frequency measures obtained from the external resources, we also included the frequency of occurrence in the textbook. Here, Count_Sum refers to the number of times a character or a word appears in the texts throughout all 12 textbooks. The textbooks contain 162,177 character tokens (Grade 1: 5445; Grade 2: 14,434; Grade 3: 25,548; Grade 4: 35,211; Grade 5: 39,223; Grade 6: 42,316) and 110,226 word tokens (Grade 1: 3,803; Grade 2: 9,863; Grade 3: 17,534; Grade 4: 23,724; Grade 5: 26,902; Grade 6: 28,400), which were derived using *Stanford CoreNLP* with Chinese model in version 4.3.2 (Manning et al., 2014). This density information reflects how frequently children are exposed to these lexical items during Chinese learning at school. Forty-eight characters, which are absent in the main texts, but present in the *Learning Activity* section, are regarded as missing values in the analysis below.

Method

Participants

One hundred and fifty adults and 66 3rd–6th graders from a Chinese elementary school were compensated for their participation in this study. All of them were native Chinese speakers. Written informed consent was obtained from all adult participants and from the parents of all children before the experiment. All participants had normal or corrected-to-normal vision and reported no reading disorder or a history of neurological or psychiatric disorder. Information of participants is shown in Table 3.

Materials

All Chinese characters and words were collected from the Appendices of the textbooks. The Appendix in each textbook includes one character list for reading, one character list for writing, and one list of words. For characters in the Reading list, children are supposed to be able to recognize them (mapping the orthographies onto their phonological forms and meanings). For those in the Writing list, children are required to write these characters correctly. Note that there is an overlap between characters in the two lists. Specifically, 2491 characters in the Writing list (which totals 2500 characters) are also present in the Reading list (which totals 3172 characters). It is often the case that a character in the Reading list will appear in the Writing list at a later grade. For example, the character "蛛" (zhū, *spider*) first appears in the Reading list in Grade 1 and then later in the Writing list in Grade 3. Such arrangement helps children consolidate their knowledge of characters and further achieve proficient reading and writing. There is no Reading list in Grade 6 and no Word

Table 3 Sample size and personal characteristics in three groups

	Younger children group		Older children group		Adult group
	Grade 3	Grade 4	Grade 5	Grade 6	
N	14	18	19	15	150
Number of females	6	9	9	8	79
Age	10.3 (.9)	11 (.6)	11.8 (.5)	13 (.5)	20.5 (2.4)
Raven's SPM score	40.5 (4.7)	43.9 (4.9)	42 (7.9)	44.1 (6.8)	55.4 (3.2)
Character recognition score	109.3 (9.3)	114.7 (11.3)	122.8 (9.3)	126.8 (6.5)	-
Word recognition score	80.9 (13.1)	92.9 (17.3)	91.6 (12.6)	103.7 (17.7)	-

Standard deviations are in parentheses. Two independent *t* tests revealed that the older children group performed significantly better than the younger children group on the character recognition task ($t(64) = – 4.25$, $p < 0.001$) and on the word reading task ($t(64) = – 1.94$, $p = 0.058$).

list in Grade 1. Therefore, our database covers all characters that are required to be recognized from Grade 1 to Grade 5 (ten volumes), characters required for writing from Grade 1 to Grade 6 (12 volumes), and all words from Grade 2 to Grade 6 (ten volumes). All words were taken into analysis regardless of their number of syllables. There are 1966 disyllabic, 82 trisyllabic, and 142 tetrasyllabic words. Only nine disyllabic words are repeated in different grades. In the word-naming task, a total number of 2999 characters (Grade 1: 700; Grade 2: 899; Grade 3: 501; Grade 4: 479; Grade 5: 399; Grade 6: 3) and 2182 words (Grade 2: 496; Grade 3: 498; Grade 4: 426; Grade 5: 374; Grade 6: 388) were used as stimuli. They were randomly split into five sub-lists for the adult group. Four of the lists consisted of 600 characters and 436 words (1036 items), and one list consisted of 599 characters and 438 words (1037 items). For child participants, the number of lists was expanded to ten in order to reduce the experimental time for each of their visits. As a result, children were required to name 518 or 519 characters and words every time they performed the task.

Procedure

The present study was approved by the Ethical Committee of the State Key Laboratory of Cognitive Neuroscience and Learning at Beijing Normal University. Prior to the word-naming task, all participants were asked to complete a non-verbal Raven's Standard Progressive Matrices (SPM, Raven et al., 1983) to ensure that they exhibit typical cognitive abilities. This test is crucial for child participants as it enables us to ascertain their typical developing status. Child participants were additionally required to complete two extra individual differences tasks: A Chinese character recognition task adopted from Li et al. (2012); Lei et al., (2011), and a word reading task (Zhang et al., 2012), both of which are used to measure children's vocabulary knowledge and reading fluency. In the Chinese character recognition task, children read out aloud 150 characters listed in the order of increasing difficulty level; the procedure stopped if they failed to recognize 15 consecutive items. In the word reading task, children were asked to read 180 disyllabic words as quickly and accurately as possible; the score was calculated by dividing the number of correct responses by the time spent on this task. Children's performance on two reading ability tasks was used as supplementary evidence to substantiate the differentiation between younger (3rd and 4th grader) and older children (5th and 6th grader) group in the analysis.

In the word-naming task, adult participants were randomly assigned to one of the five sub-lists, with each participant only naming one list. Each child participant completed three individual difference tasks on the first day and the naming task in the following few days. On average, each 3rd and 4th grader completed four or five sub-lists, while each 5th and 6th grader completed five or six sub-lists on separate days. As a result, each character or word received 30 responses from adult participants, 15 responses from children in Grade 3 and 4, and 20 responses from children in Grade 5 and 6. During the experiment, each sub-list was further divided into ten blocks with an equal number of trials. Participants were allowed to take a short break between every two consecutive blocks. All stimuli were presented in black with the *SimSun* font of size 30 at the center of a white screen. Participants sat in front of a computer at a distance of approximately 80 cm from the screen. Each trial began with a fixation presented at the center of the screen when participants were required to press the space bar to continue. Then, a target character or word appeared after a blank screen of 300 ms. Participants were asked to name the character or word as quickly and accurately as possible. The target disappeared upon a naming response or after 3000 ms without a naming response. Participants took a practice session with 20 trials, which were not included in experimental lists, before the formal experiment. The entire experiment lasted 30–40 minutes.

Data analyses

Naming accuracy (ACC) was coded manually by the experimenter. Incorrect responses (accounting for 2.53% of all responses) and absolute outliers (RTs that were faster than 200 ms or slower than 2500 ms) were excluded from further analysis. All participants achieved an overall ACC higher than 64%. We also rejected RTs over 2.5 standard deviations away from the mean for every participant. These trimming steps led to an exclusion of 5.95% of the data. To minimize individual differences in timing, we transformed each participant's RTs into z scores. The average z scores across participants yielded zRTs for each item. ACC across participants for each item was also calculated. Numerical lexical variables, including the number of strokes, number of radicals, number of meanings, number of pronunciations, and Count_Sum, were log-transformed to better approximate a normal distribution as they are highly positively skewed (Hair et al., 2009). Categorical variables (i.e., Structure and Regularity) with three levels were independently coded into two new variables based on the dummy coding system (Hair et al., 2009) in line with previous studies (Chang et al., 2016; Liu et al., 2007; Wang et al., 2020). The structure was divided into a left-right structure and top-down structure. For the left-right structure variable, characters with a left-right structured were coded as 1, while others were coded as 0. The same criteria were applied to the top-down structure variable, with top-down structured characters coded as 1 and others coded as 0. Regularity was separated into regular and semi-regular variables, and a similar coding system was employed. For the regular variable, characters with regular pronunciation were coded as 1, and others were coded as 0. For the semi-regular variable, characters with semi-regular pronunciations were coded as 1, and all others were coded as 0.

The goal of our data analyses was threefold. First, to trace the gradual and subtle development in characters (in both Reading and Writing list) and words that children need to learn at elementary schools, we attempt to capture potential differences in those variables that may reflect essential features across six grades through Poisson regression and simple regression analyses. Due to the fact that most variables were count data, such as nStroke, nRadical, nPronunciation, nMeaning, Count_Sum, the Poisson regression analyses were performed to model these measures as a function of Grade; to compensate for the overdispersion identified in some variables, negative binomial models were used instead (Ismail & Jemain, 2007). Simple regression was used for the continuous variable logCHR-CD.

Second, to identify significant predictors for naming RTs and ACC among these lexical variables and investigate the relative importance of those predictors, we used a stepwise multiple regression analysis (MRA). For the word-level analysis, we included word-level variables (volume, length, nPronunciation, nMeaning, Count_Sum and logW-CD), character-level variables of the first (C1) and the second (C2) character of the word (lexical variables of the third (C3) and forth (C4) character were not included because three- and four-character words accounted for only 10% of the current database), and the sum value of C1-C4 (the log-transformed values across characters were summed, Sum_nStroke, Sum_nRadical, Sum_nPronunciation, Sum_nMeaning, Sum_Count-Sum, and Sum_logCHR_CD; see Table 9). It should be pointed out that averaged character-level factors were used in word-level analysis in previous studies such as Tsang et al., (2018). However, we used the summed value on the character-level to better capture the overall features of words. Items with zRTs longer or shorter than 2.5 standard deviations from the mean, as well as items with low ACC (< 50%) were removed before running the MRA (Sun et al., 2018). The application of the criterion resulted in the inclusion of 2677 characters and 2146 words in the younger children group, 2881 characters and 2182 words in the older children group, and 2941 characters and 2164 words in the adults group. All of the reported models below were constructed using a stepwise multiple regression and excluded cases listwise. Furthermore, the relative importance analysis was used to reveal how each variable contributes to the naming task performance. We calculated the *lmg* metric for variables entered the final model using the *relaimpo* package (Grömping, 2006) to examine the relative contribution of any particular variable.

Third, to understand how lexical effects would vary across three different aged-groups, we conducted a between-group analysis to compare the magnitude of those effects. Considering the abundance of predictors in current study, we selected those predictors which were revealed to be significant in all three groups in all analyses (naming RTs and ACC, on the character and word level) to determine the trajectory of reading performance among different age groups.

Results

Responses were divided into three groups: the younger children group, the older children group, and the adult group. The younger children group (mean age was 10.65, mean Raven's SPM score was 42.2, mean character and word recognition scores were 112 and 86.9, respectively) included 3rd and 4th graders, and the older children group (mean age was 12.4, mean Raven's SPM score was 43.05, mean character and word recognition scores were 124.8 and 97.65, respectively) included 5th and 6th graders. The result of independent *t* tests showed that the reading performance in the older children group was significantly better than that in the younger children group (see the Note in Table 3).

The character level

Table 4 shows the descriptive statistics of lexical variables and naming responses for all Chinese characters. Note that the frequency-related variables (logCHR-CD) of seven characters were not found in Cai and Brysbaert (2010). Table 5 presents the correlation (Pearson's *r*) matrix among the log-transformed numerical variables and the character naming performance for the three groups. It shows that naming latency and accuracy for of all the three groups were highly correlated to all the variables except for the number of pronunciations.

Differences across the six grade levels

In the Reading list (see Fig. 1), it is obvious that nStroke and nRadical increase with grade level. Difference on nStroke is not attested in the comparison between Grades 2 and 3 ($p = 0.51$) and between Grades 4 and 5 ($p = 0.16$) according to the Tukey's tests. nRadical in Grade 1 is significantly smaller than other grade levels. nMeaning, on the other hand, decreases with grades, but further analysis reveals no significant difference between Grades 2 and 3 ($p = 0.12$) and

Table 4 Descriptive statistics of 2999 characters

	N	Min	Max	Mean
RT_younger children	2951	315	1805	835 (151)
RT_older children	2994	522	1710	761 (127)
RT_adult	2999	511	1124	664 (81)
ACC_younger children	2999	0	1	0.85 (0.24)
ACC_older children	2999	0	1	0.92 (0.16)
ACC_adult	2999	.10	1.00	0.97 (0.09)
nStroke	2999	1	23	9.56 (3.39)
nRadical	2999	1	7	2.74 (1.03)
nPronunciation	2999	1	5	1.22 (0.50)
nMeaning	2999	1	18	3.00 (2.14)
Count_Sum	2951	1	6498	57.06 (215.61)
LogCHR-CD	2992	.00	3.80	2.88 (0.74)
LogCHR-CD-C	2998	2.78	7.6	4.93 (0.68)

RT, reaction time; ACC, accuracy rate; nStroke, number of strokes; nRadical, number of radicals; nPronunciation, number of pronunciations; nMeaning, number of meanings. Count_Sum; total number of times the character occurs in all texts across six grades; LogCHR-CD, logarithmic contextual diversity (CD) (based on SUBTLEX-CH, Cai & Brysbaert, 2011); LogCHR-CD-C, logarithmic CD which is based on children's frequency CCLLOWW (Li et al., 2022).

between Grades 4 and 5 ($p = 0.06$). There is no significant difference in nPronunciation across grades. Comparisons across all grades on Count_Sum and logCHR-CD indicate that these two frequency-related variables significantly decrease with grades (all *p*s < .05).

Similar to the Reading list (see Fig. 2), differences in nPronunciation across the six grades in the Writing list are not significant (all *p*s > 0.77). nStroke shows an increasing trend as grade advances, but exceptions are found in the comparison between Grades 3 and 5 and between Grades 4 and 5 (*p*s > 0.09). Differences across grades in nRadical converge on the comparison between Grade 1 and other grades (all *p*s < .01). Grades significantly differ on nMeaning and Count_Sum, except that the differences between Grade 5 and Grade 6 on nMeaning ($p = 0.64$) and on Count_Sum ($p = 0.05$) are not significant.

The effects of the lexical variables

In the multiple regression analysis for naming zRTs and ACC, the numerical variables, Volume, logCHR-CD, Count_Sum, nStrokes, nRadicals, nMeaning, and nPronunciation were entered in Block 1. Next, categorical variables, character phonograms, regularity (regular, semi-regular) and Structure (left-right, top-down) were entered in Block 2.

For the naming zRTs, the result showed that Volume, logCHR-CD, Count_Sum, and nStroke in Block 1, and left-right structure, top-down structure, and regular in Block 2 had significant effects on zRTs in the younger children group (overall $R^2 = .351$, F (7, 2894) = 223.89, $p < 0.001$) (see Table 6). In the older children group, the effects of Volume, logCHR-CD, Count_Sum, nMeaning, and nPronunciation in Block 1, and left-right structure, top-down structure, phonograms and regular in Block 2 were significant ($R^2 = .443$, F (9, 2930) = 248.95, $p < 0.001$), see Table 7 for details. For the adult group, results showed that Volume, logCHR-CD, nMeaning, and nPronunciation in Block 1 and left-right structure and semi-regular in Block 2 were significant contributors to the naming zRTs ($R^2 = .384$, F (7, 2929) = 256.14, $p < 0.001$). The effect of nStroke was

Table 5 Correlations of lexical variables and naming performance (zRT and ACC) for characters

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. zRT_younger children	1												
2. zRT_older children	$.715^{**}$	1											
3. zRT_adult	$.604^{**}$	$.758^{**}$	1										
4. ACC_younger children	$-.578^{**}$	$-.695^{**}$	$-.631^{**}$	1									
5. ACC_older children	$-.470^{**}$	$-.641^{**}$	$-.624^{**}$	$.832^{**}$	1								
6. ACC_adult	$-.357^{**}$	$-.529^{**}$	$-.627^{**}$	$.615^{**}$	$.754^{**}$	1							
7. Volume	$.460^{**}$	$.478^{**}$	$.422^{**}$	$-.522^{**}$	$-.382^{**}$	$-.286^{**}$	1						
8. nStroke	$.250^{**}$	$.247^{**}$	$.235^{**}$	$-.179^{**}$	$-.142^{**}$	$-.126^{**}$	$.287^{**}$	1					
9. nRadical	$.198^{**}$	$.201^{**}$	$.186^{**}$	$-.149^{**}$	$-.117^{**}$	$-.096^{**}$	$.264^{**}$	$.727^{**}$	1				
10. nPronunciation	$-.046^{*}$	– 0.006	$.043^{*}$	0.034	0.019	– 0.030	$-.054^{**}$	$-.053^{**}$	– 0.017	1			
11. nMeaning	$-.333^{**}$	$-.367^{**}$	$-.337^{**}$	$.296^{**}$	$.248^{**}$	$.216^{**}$	$-.338^{**}$	$-.266^{**}$	$-.210^{**}$	$.338^{**}$	1		
12. Count_Sum	$-.503^{**}$	$-.518^{**}$	$-.448^{**}$	$.500^{**}$	$.403^{**}$	$.294^{**}$	$-.725^{**}$	$-.343^{**}$	$-.303^{**}$	$.117^{**}$	$.465^{**}$	1	
13. logCHR-CD	$-.528^{**}$	$-.615^{**}$	$-.602^{**}$	$.568^{**}$	$.492^{**}$	$.427^{**}$	$-.527^{**}$	$-.312^{**}$	$-.251^{**}$	$.100^{**}$	$.522^{**}$	$.647^{**}$	1

nStroke, number of strokes; nRadical, number of radicals; nPronunciation, number of pronunciations; nMeaning, number of meanings, RT, reaction time; ACC, accuracy rate. * Correlation is significant at the 0.05 level (two-tailed). ** Correlation is significant at the 0.01 level (two-tailed). All variables are log-transformed except for the ACC, zRT and Volume.

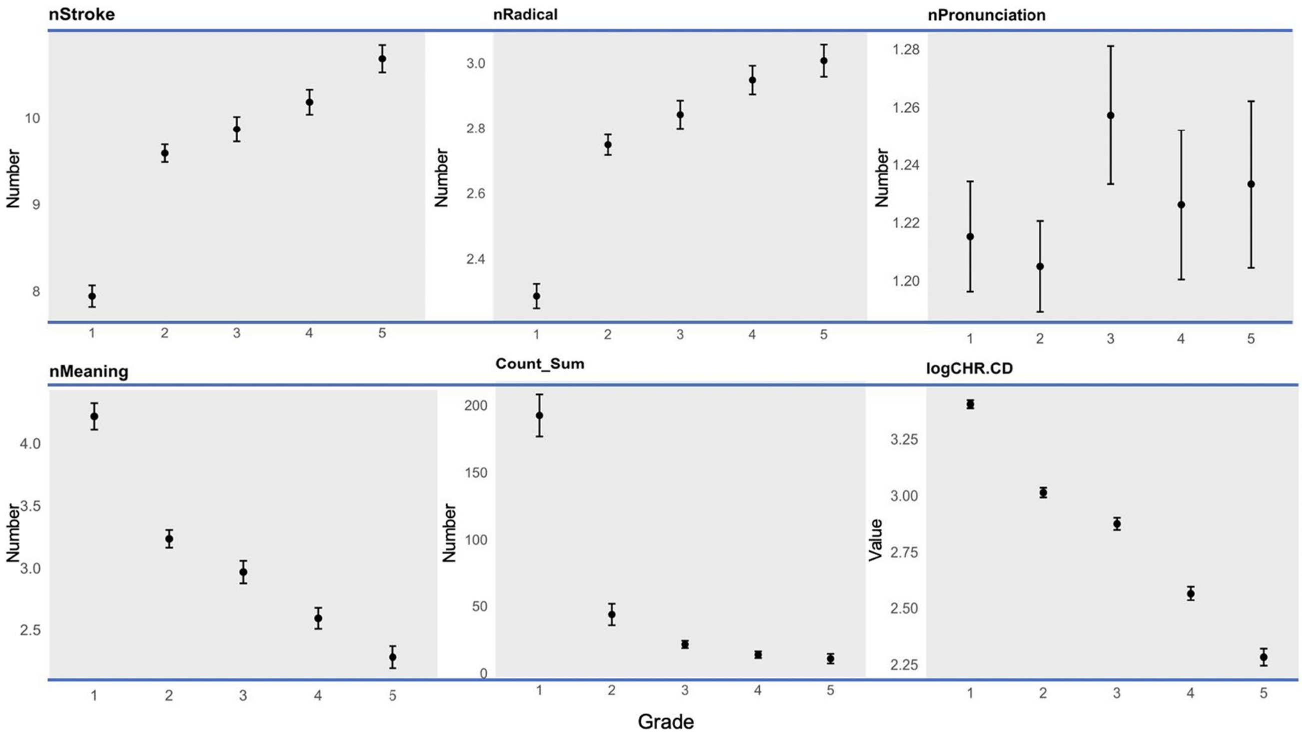

Fig. 1 Critical predictors across five grades in the Reading list in Chinese character database. The *error bar* represents standard errors. Variables such as nStroke, nRadical, nPronunciation, nMeaning, and Count_Sum are counted as numbers; logCHR-CD is log-transformed

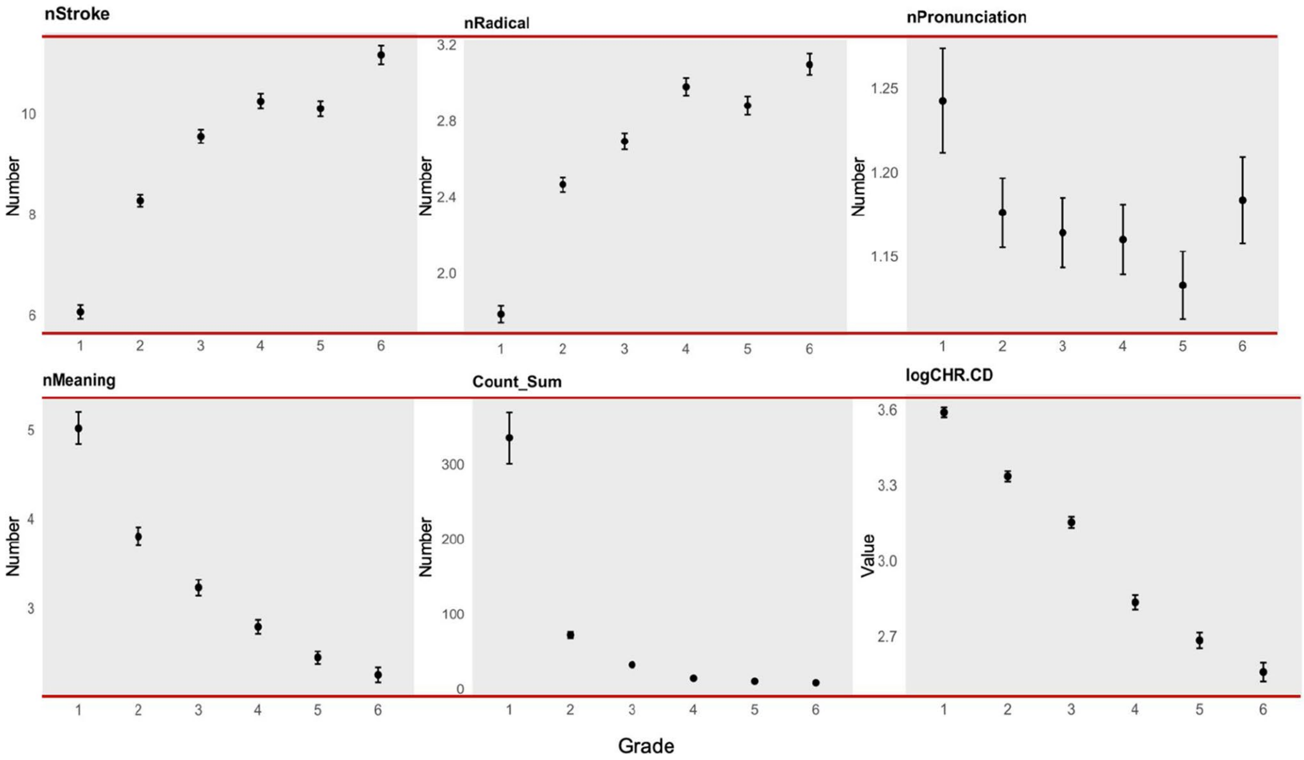

Fig. 2 Critical predictors across five grades in the Writing list in Chinese character database. The *error bar* represents standard errors. Variables such as nStroke, nRadical, nPronunciation, nMeaning, and Count_Sum are counted as numbers; logCHR.CD is log-transformed

Table 6 Results of stepwise MRA on zRT and ACC in the younger children group on the character level

	B	*t*	*p*	*ΔR*^{2}	*lmg*
zRT					
Intercept	0.605	10.048	0.000		
Volume	0.025	7.382	0.000	0.048	33.10
logCHR-CD	– 0.195	– 16.565	0.000	0.279	51.41
Count_Sum	– 0.101	– 6.510	0.000	0.011	2.22
nStroke	0.143	3.245	0.001	0.002	6.78
left-right	0.041	2.093	0.036	0.001	4.00
top-down	– 0.057	– 2.661	0.008	0.007	1.52
regular	– 0.053	– 3.639	0.000	0.003	0.96
R^2 = .351					
ACC					
Intercept	0.528	17.166	0.000		
Volume	– 0.024	– 13.752	0.000	0.069	37.92
logCHR-CD	0.133	21.204	0.000	0.324	49.12
Count_Sum	0.020	2.430	0.015	0.001	1.31
nStroke	0.019	0.890	0.374	0.001	2.32
nMeaning	– 0.017	– 1.270	0.204	0.001	6.92
top-down	0.017	2.241	0.025	0.001	0.32
regular	0.063	8.390	0.000	0.014	2.09
R^2 = .410					

nStroke, number of strokes; nMeaning, number of meanings, left-right, left-right structure; top-down, top-down structure. *lmg* Measures (relative importance) are scaled to 100%.

Table 7 Results of stepwise MRA on zRT and ACC in the older children group on the character level

	B	*t*	*p*	*ΔR*^{2}	*lmg*
zRT					
Intercept	0.553	15.558	0.000		
Volume	0.018	6.963	0.000	0.031	24.16
logCHR-CD	– 0.217	– 22.974	0.000	0.379	54.04
Count_Sum	– 0.055	– 4.520	0.000	0.005	1.60
nMeaning	– 0.063	– 2.886	0.004	0.002	11.03
nPronunciation	0.226	5.055	0.000	0.004	0.96
left-right	0.036	2.368	0.018	0.007	3.13
top-down	– 0.038	– 2.397	0.017	0.001	1.02
phonograms	0.032	2.322	0.020	0.001	3.20
regular	– 0.056	– 4.752	0.000	0.004	0.87
R^2 = .443					
ACC					
Intercept	0.666	28.860	0.000		
Volume	– 0.008	– 5.985	0.000	0.020	28.73
logCHR-CD	0.086	19.145	0.000	0.243	65.24
Count_Sum	0.016	2.584	0.010	0.001	1.34
nStroke	0.019	1.186	0.236	0.001	2.57
nPronunciation	– 0.047	– 2.214	0.027	0.001	0.18
top-down	0.013	2.270	0.023	0.001	0.70
regular	0.028	5.045	0.000	0.006	1.23
R^2 = .274					

nStroke, number of strokes; nPronunciation, number of pronunciations; nMeaning, number of meanings. *lmg* Measures (relative importance) are scaled to 100%.

significant ($p = 0.02$) in Block 1 before the factors in Block 2 were entered, see Table 8 for details. Given that Volume, logCHR-CD, and left-right structure were significant predictors in all three groups, we analyzed how the effects of these predictors might vary across groups. Results showed that the magnitudes of Volumn and logCHR-CD effects in the younger and older children groups were significantly greater than that in the adult group (*p*s < 0.001); however, the differences between the younger children and the older children group were not significant (*p*s = 0.188). For the left-right structure, group comparisons only revealed a significant difference between the younger children group and the adult group ($p < 0.001$). Details are shown in Table 9.

For ACC, we found that Volume, logCHR-CD, Count_Sum, nStroke and nMeaning in Block1, and top-down structure and regular in Block 2 were entered into the final model ($R^2 = .410$, *F* (7, 2936) = 291.17, $p < 0.001$) (see Table 6) in the younger children group. In the older children group, Volume, logCHR-CD, Count_Sum, nStroke and nPronunciation in Block 1, and top-down structure and regular in Block 2 was found to be significant ($R^2 = .274$, *F* (7, 2936) = 158.32, $p < 0.001$) (see Table 7). In the adult group, logCHR-CD, Volume and nPronunciation in Block 1 were entered into the final model ($R^2 = .161$, *F* (3, 2933) = 188.65, $p < 0.001$) (see Table 8). The comparisons between three groups on the effects of Volume and logCHR-CD showed a gradient trend, with the younger children group showing the largest magnitude, followed by the older children group, and finally the adult group with the lowest magnitude (*p*s < 0.001). See Table 10 for details.

Results of relative importance in this analysis showed that logCHR-CD and Volume are the two strongest predictors for both naming latencies and accuracy across three groups. In addition, variance inflation factors (VIFs) were less than 3 for all variables that were added into the final models, excluding the possibility of collinearity (Hair, 2011; O'Brien, 2007). As noted by Plonsky & Ghanbar (2018), R^2 equal to or less than 0.2 indicates a small effect size. All models in the current analysis provided good fits to the character naming data except for the ACC in the adult group (R^2 = 0.161).

The word level

Table 11 shows the descriptive statistics of a total of 2182 Chinese words on lexical variables and responses from the

Table 8 Results of stepwise MRA on zRT and ACC in the adult group on the character level

	B	*t*	*p*	ΔR^2	*lmg*
zRT					
Intercept	0.278	9.578	0.000		
Volume	0.010	7.083	0.000	0.015	48.3
logCHR-CD	– 0.147	– 24.927	0.000	0.339	56.7
nStroke	0.020	0.994	0.320	0.001	4.1
nMeaning	– 0.044	– 3.163	0.002	0.004	11.0
nPronunciation	0.237	8.162	0.000	0.013	3.3
left-right	0.046	6.698	0.000	0.010	5.6
semi-regular	0.022	2.505	0.012	0.001	0.9
$R^2 = .384$					
ACC					
Intercept	0.897	136.993	0.000		
Volume	– 0.002	– 3.928	0.000	0.004	21.9
logCHR-CD	0.031	17.644	0.000	0.147	73.7
nPronunciation	– 0.052	– 5.768	0.000	0.010	4.3
$R^2 = .161$					

nStroke, number of strokes; nPronunciation, number of pronunciations; nMeaning, number of meanings, *lmg* Measures (relative importance) are scaled to 100%.

word-naming task. Note that the frequency-related variables (logW-CD) of 110 characters were not found in the database of Cai and Brysbaert (2010). Table 12 demonstrates the correlation (Pearson's *r*) matrix among the log-transformed numerical variables and the word-naming performance (zRT and ACC) in three groups. It shows that zRT in three groups were highly correlated with all the word-level variables, character-level variables of C1 and C2, and total number of strokes and radicals of the whole-word except the number of pronunciations. In addition, naming accuracy was highly correlated with Volume, Count_Sum and logW-CD at the word level, and logCHR-CD, Count_Sum of C1, C2 at the character level.

Differences across the six grade levels

As shown in Fig. 3, the comparisons across all grades at the word level reveal no significant differences on Length or nPronunciation. This pattern is also shown in nMeaning, except that the comparison between Grade 2 and Grade 5 is significant ($p < 0.05$). Sum_nStroke increases as the grade progresses, but this trend is not statistically attested in the comparison between Grade 5 and Grade 6 ($p = 0.88$). Analysis of Count_Sum reveals that comparisons between any grades are significant (*p*s < 0.05). Grades differ in logW-CD except for the comparisons between Grade 3 and 4 ($p = 0.64$) and between Grade 5 and 6 ($p = 0.78$).

The effects of the lexical variables

For naming zRTs, we found that logW-CD, Volume, Sum_nStroke, and C1_Count_Sum were strong predictors in the younger children group, along with other variables ($R^2 = .304$, F (8, 2059) = 112.27, $p < 0.001$), see Table 13 for details. In the older children group, strong predictors were logW-CD, C1_CS, C1_logCHR-CD and Sum_nStr; the final model accounts for 32.8% of the variance (F (9, 2062) = 112.03, $p < 0.001$) (see Table 14). In the adult group, logW-CD and C1_logCHD-CD, along with other variables, were found to significantly affect the naming RTs ($R^2 = .206$, F (9, 2051) = 59.17, $p < 0.001$) (see Table 15). Among these predictors, Volumn, logW-CD, C1_Count_Sum, C2_Count_Sum, and Sum_nStrokes were shown to significantly affect the naming RTs in all three groups. Results from the group comparisons indicate that the effects of Volume, C2_Count_Sum, and Sum_nStrokes are particularly pronounced in the younger children group

Table 9 Results of regression coefficient comparisons on zRT among three groups on the character level

Variables	Group 1	Group 2	b1	b2	difference	*t*	*p*
Volume	Younger children	Older children	0.039	0.026	0.013	2.191	0.028
	Younger children	Adult	0.039	0.012	0.027	10.063	< 0.001
	Older children	Adult	0.026	0.012	0.014	– 5.567	< 0.001
logCHR-CD	Younger children	Older children	– 0.228	– 0.238	0.01	1.316	0.188
	Younger children	Adult	– 0.228	– 0.172	– 0.056	6.377	< 0.001
	Older children	Adult	– 0.238	– 0.172	– 0.066	– 7.997	< 0.001
left-right structure	Younger children	Older children	0.086	0.07	0.016	0.985	0.325
	Younger children	Adult	0.086	0.054	0.033	4.299	< 0.001
	Older children	Adult	0.07	0.054	0.016	0.933	0.351

Table 10 Results of regression coefficient comparisons on ACC among three groups on the character level

Variables	Group 1	Group 2	b1	b2	difference	*t*	*p*
Volume	Younger children	Older children	– 0.026	– 0.01	– 0.016	– 10.667	< 0.001
	Younger children	Adult	– 0.026	– 0.002	– 0.024	7.334	< 0.001
	Older children	Adult	– 0.01	– 0.002	– 0.007	– 17.584	< 0.001
logCHR-CD	Younger children	Older children	0.129	0.087	0.041	16.87	< 0.001
	Younger children	Adult	0.129	0.046	0.083	– 6.71	< 0.001
	Older children	Adult	0.087	0.046	0.042	10.614	< 0.001

than other two groups. The two child groups also exhibited greater logW-CD and C1_Count_Sum effects than the adult group, though there was no significant difference on logW-CD between the two child groups ($p = 0.065$). See Table 16 for details.

On the other hand, we found that Volume, logW-CD, C1_logCHR-CD, and C1_Count_Sum and C2_Count_Sum exerted significant effects on naming ACC for the younger children group. Along with other variables, the final model for this group accounts for 15.9% of the variance (F (8, 2063) = 48.66, $p < 0.001$) (see Table 13). For the older children group, it should be noted that Sum_nPro also made significant contribution to account for the variance except for logW-CD and Volume ($R^2 = .090$, F (6, 2065) = 33.95, $p < 0.001$) (see Table 14). For the adult group, the result showed that Sum_nPro, logW-CD, C1_CS, and C2_logCHR-CD were significant contributors to the naming accuracy ($R^2 = .050$, F (4, 2056) = 27.27, $p < 0.001$) (see Table 15). The comparisons between three groups on the effects of logW-CD and C2_Count_Sum showed a gradient trend, with the younger children group showing the largest magnitude, followed by the older children group, and finally the adult group with the lowest magnitude ($ps < 0.05$). See Table 17 for details.

Results of relative importance in this analysis showed that logW-CD is the strongest predictor for both word-naming latencies and accuracy across three groups. The VIF results for all variables were less than 4, excluding strong collinear relationships between them. In addition, naming accuracy for words was lower than that for characters in all three groups. All models provided good fits to the zRTs ($R^2 >$ 0.2), but not the word-naming accuracies ($R^2 < 0.2$).

Discussion

In this study, we present a large-scale database of Chinese characters and words based on the contemporary elementary school textbooks issued by the Ministry of Education in mainland China. The current database incorporates key lexical variables collected from 2999 Chinese characters and 2182 words, as well as norming data in a word-naming task from both school-age children and adults. We found that, first, visual complexity of characters and words learned in elementary schools increases with grades, whereas the semantic richness and frequency tend to decrease. Second, frequency, visual complexity, semantic and phonological attributes, structural type, and phonetic regularity are significant contributors in lexical processing. However, the effects of these predictors in lexical processing vary across children and adults. Third, the factor of age of acquisition, objectively indexed by the Volume in this study, had a significant effect in the naming task, but different patterns were attested in character and word naming.

Changes of lexical variables across six grades

The distribution of character-level variables in the Reading list from Grade 1 to Grade 6 indicates that visual complexity of character gradually increases with grade level, whereas semantic richness, frequency, and naming accuracy decrease with grade level. Similar patterns can be found in the Writing list, given that there is a considerable overlap between the two lists. With regard to the development of variables in the word list, it can be clearly seen that three or four-characters are clustered in later grades, which explains the increase of visual complexity across grades. Another notable trend is that the percentage of polysemes (i.e., words with multiple lexical meanings) and high-frequency words decrease steadily. All these findings reflect a fundamental principle in literacy acquisition: characters and words to be learned tend to become more structurally complicated, less frequent, and semantically unambiguous.

These findings are consistent with those of Shu et al. (2003). Moreover, we expanded Shu et al., (2003)'s database in several aspects. First, the current database explicitly indicates the requirement (i.e., reading or writing) for characters learned at a certain level, which is referenced in the Reading and Writing lists in the new edition of the textbooks. Second, the current database covers a wider

Table 11 Descriptive statistics of 2182 words

	N	Min	Max	Mean (SD)
RT_yonger children	2177	554	1406	775 (111)
RTs_older children	2182	518	1067	680 (68)
RTs_adult	2182	516	870	623 (45)
ACC_younger children	2182	0.00	1.00	0.94 (0.15)
ACC_older children	2182	0.25	1.00	0.97 (0.08)
ACC_adult	2182	0.27	1.00	0.99 (0.04)
Length	2182	2	4	2.17 (0.52)
nPronunciation	2182	1	2	1.01 (0.12)
nMeaning	2182	1	6	1.46 (0.78)
Count_Sum	2182	1	238	6.70 (14.56)
logW-CD	2072	0.00	3.80	2.06 (0.87)
C1_nStroke	2182	1	21	8.41 (3.30)
C2_nStroke	2182	2	21	8.33 (3.20)
C3_nStroke	225	1	18	7.12 (3.35)
C4_nStroke	142	2	17	8.43 (3.33)
Sum_nStroke	2182	5	54	18.03 (6.06)
C1_nRadical	2182	1	6	2.45 (1.03)
C2_nRadical	2182	1	6	2.38 (0.98)
C3_nRadical	225	1	6	2.06 (1.05)
C4_nRadical	142	1	6	2.41 (1.00)
Sum_nRadical	2182	2	17	5.20 (1.79)
C1_nPronunciation	2182	1	5	1.18 (0.44)
C2_nPronunciation	2182	1	6	1.20 (0.50)
C3_nPronunciation	225	1	6	1.26 (0.60)
C4_nPronunciation	142	1	5	1.25 (0.58)
Sum_nPronunciation	2182	2	13	2.59 (0.99)
C1_nMeaning	2182	1	18	4.36 (2.74)
C2_nMeaning	2182	1	18	4.59 (2.95)
C3_nMeaning	225	1	14	4.65 (2.99)
C4_nMeaning	142	1	14	4.35 (2.58)
Sum_nMeaning	2182	2	39	9.71 (5.01)
C1_Count_Sum	2182	1	3875	168.95 (418.09)
C2_Count_Sum	2181	1	6498	193.22 (371.65)
C3_Count_Sum	225	3	3875	337.53 (640.56)
C4_Count_Sum	142	1	1322	152.39 (239.18)
Sum_Count_Sum	2181	2	8151	406.98 (676.05)
C1_logCHR-CD	2182	0.00	3.80	3.36 (0.50)
C2_logCHR-CD	2182	0.00	3.80	3.39 (0.51)
C3_logCHR-CD	225	1.20	3.80	3.42 (0.51)
C4_logCHR-CD	142	1.20	3.80	3.43 (0.51)
Sum_logCHR-CD	2182	2.64	15.17	7.32 (2.04)
C1_logCHR-CD-C	2182	1.52	3.33	2.84 (0.26)
C2_logCHR-CD-C	2182	0.90	3.33	2.87 (0.27)
C2_logCHR-CD-C	225	2.03	3.33	2.92 (0.27)
C2_logCHR-CD-C	142	1.99	3.30	2.88 (0.26)
Sum_logCHR-CD-C	2182	3.69	12.93	2.60 (1.62)

RT, reaction time; ACC, accuracy rate; C1-C4, the first to forth character of the word; nStroke, number of strokes; nRadical, number of radicals; nPronunciation, number of pronunciations; nMeaning, number of meanings; Count_Sum; total number of times the character occurs in all texts across six grades; logCHR-CD, logarithmic contextual diversity (CD) (based on SUBTLEX-CH, Cai & Brysbaert, 2011); logCHR-CD-C, logarithmic CD which is based on children's frequency CCLLOWW (Li et al., 2022)

range of words (up to 2182 words) than Shu et al. (2003), in light of the fact that both characters and words are building blocks of the Chinese lexicon, and most language units or concepts are represented in words as they can convey relatively complete meanings. Third, we collected naming data from both children at different grades and adults, which allows us to examine the developmental trajectory of the effects of lexical variables in visual word processing. Therefore, the word cluster in the current database arranged by grade provides a useful resource to study literacy acquisition among school-aged children.

Influential factors for character and word processing

We found that factors such as frequency, volume, visual complexity, structural type, semantic and pronunciation richness, and phonetic regularity contributed significantly to naming latency and accuracy for both characters and words in three groups of participants. Differences between children and adults were also observed in the current study. However, it should be noted that regression models did not fit well for character naming accuracy in the adult group and word-naming accuracy in all three groups. We speculated that this could be due to relatively high naming accuracy and less variances for items with different attributes (see Table 4 for character and Table 9 for word). As a result, the overall models for accuracy failed to fit well, and thus we mainly discussed the results of naming latency. These results were consistent with previous databases based on adult materials (Liu et al., 2007; Sun et al., 2018; Sze et al., 2014; Tsang et al., 2018). Firstly, character frequency retrieved from Cai and Brysbaert (2010) was found to be the strongest predictor in both naming latency and accuracy across three groups. Specifically, words composed of more frequent characters led to faster responses and higher accuracy, compared with words composed of less frequent characters. Consistent with the results of Sun et al. (2018) and Li et al. (2022), this finding demonstrates significant frequency effect for both the first and the second characters in the word-naming task. Moreover, the frequency effect was notably stronger in the younger children group than in the other two groups, reflecting a significant impact of frequency on lexical processing at an early stage of reading acquisition. However, the number of times that characters occur in the textbooks (Count_Sum) only showed a significant effect on character naming latency and accuracy in children but not in adults. In the word level, Count_Sum of the constituting characters, mostly the first two characters, made a significant contribution to word processing in all three groups, but accounted for less variance in adults than in children. This is probably due to the effect of recent exposure (Kaschak, 2007).

Table 12 Correlations coefficient matrix of variables for words

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. zRT_younger children	1														
2. zRT_older children	.688**	1													
3. zRT_adult	.470**	.638**	1												
4. ACC_younger children	– .422**	– .441**	– .340**	1											
5. ACC_older children	– .220**	– .308**	– .274**	.652**	1										
6. ACC_adult	– .194**	– .267**	– .283**	.458**	.601**	1									
7. Volume	.364**	.272**	.212**	– .281**	– .134**	– .095**	1								
8. Length	.248**	.213**	.193**	– .079**	0.000	– 0.016	.086**	1							
9. Count_Sum	– .376**	– .353**	– .255**	.222**	.118**	.090**	– .420**	– .224**	1						
10. logW-CD	– .392**	– .429**	– .369**	.247**	.175**	.151**	– .199**	– .255**	.527**	1					
11. nPronunciation	– .045*	– .056**	– 0.032	0.027	0.013	– 0.002	– .045*	– 0.037	.090**	.058**	1				
12. nMeaning	– .158**	– .180**	– .135**	.094**	.063**	.061**	– .052*	– .180**	.203**	.283**	.248**	1			
13. C1_nstroke	.127**	.186**	.120**	– .047*	– 0.041	– 0.032	.183**	– .158**	– .077**	– .051*	-0.010	-0.018	1		
14. C1_nRadical	.119**	.180**	.121**	– .056**	– .063**	– 0.031	.182**	– .156**	– .069**	– .044*	-0.005	-0.004	.741**	1	
15. C1_ nPronunciation	– .054*	– 0.025	– 0.017	– .082**	– .138**	– .087**	– 0.009	0.001	0.018	.045*	.063**	0.014	– .050*	-0.015	1
16. C1_nMeaning	– .164**	– .233**	– .216**	.084**	0.033	.061**	– .108**	.052*	0.031	.136**	0.025	.104**	– .296**	– .234**	.319**
17. C1_Count_Sum	– .272**	– .353**	– .240**	.214**	.134**	.087**	– .330**	.161**	.273**	.138**	0.039	.048*	– .470**	– .419**	.141**
18. C1_logCHR-CD	– .254**	– .376**	– .352**	.239**	.161**	.156**	– .191**	.090**	.125**	.358**	0.019	.106**	– .313**	– .284**	.129**
19. C2_nstroke	.154**	.125**	.101**	– .114**	– .081**	– 0.035	.220**	– 0.037	– .104**	– .065**	-0.029	-0.028	0.025	0.004	0.021
20. C2_nRadical	.140**	.116**	.094**	– .088**	– .071**	– .046*	.205**	– .066**	– .083**	– .058**	-0.039	-0.025	0.024	0.028	0.027
21. C2_ nPronunciation	.052*	.047*	0.030	– .069**	– .120**	– .091**	0.022	– 0.008	0.014	0.023	0.035	0.033	0.022	– 0.011	.083**
22. C2_nMeaning	– .106**	– .106**	– .077**	.099**	0.031	.053*	– .147**	.047*	.090**	.113**	.065**	.128**	– 0.025	– 0.011	0.025
23. C2_Count_Sum	– .235**	– .199**	– .143**	.188**	.116**	.078**	– .325**	.073**	.316**	.164**	.091**	.065**	– 0.004	– 0.017	-0.018
24. C2_logCHR-CD	– .197**	– .219**	– .208**	.201**	.136**	.148**	– .207**	.043*	.156**	.337**	.069**	.107**	– 0.030	– 0.010	-0.005
25. Sum_nStroke	.338**	.318**	.261**	– .142**	– .050*	– 0.038	.248**	.803**	– .270**	– .272**	– .047*	– .177**	.287**	.168**	– 0.017
26. Sum_nRadical	.282**	.248**	.207**	– .160**	– .092**	– .045*	.393**	.240**	– .224**	– .186**	– .043*	– .069**	.341**	.452**	– 0.007
27. Sum_nPronunciation	0.039	.054*	0.041	– .106**	– .168**	– .122**	0.024	.173**	– 0.016	– 0.002	.055*	– 0.005	– .048*	– .050*	.660**
28. Sum_nMeaning	– 0.013	– .068**	– .054*	.058**	0.039	.050*	– .093**	.590**	– .058**	0.014	0.028	0.026	– .258**	– .215**	.177**
29. Sum_ Count_Sum	– .073**	– .115**	– .049*	.134**	.113**	.063**	– .238**	.741**	.118**	– 0.012	0.033	– .067**	– .314**	– .293**	.056**
30. Sum_logCHR_CD	.102**	0.032	0.026	.048*	.079**	.064**	– 0.027	.912**	– .125**	– 0.031	– 0.010	– .105**	– .229**	– .217**	0.033
	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
16. C1_nMeaning	1														
17. C1_Count_Sum	.506**	1													
18. C1_logCHR-CD	.550**	.656**	1												
19. C2_nstroke	0.031	– 0.023	– 0.006	1											

Table 12 (continued)

	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
20. C2_nRadical	0.014	− 0.012	− 0.003	.736**	1										
21. C2_ nPronunciation	0.002	− 0.025	− 0.003	− 0.018	0.038	1									
22. C2_nMeaning	.058**	0.014	.051*	− .347**	− .325**	.290**	1								
23. C2_Count_Sum	− 0.021	.085**	0.007	− .439**	− .401**	.094**	.558**	1							
24. C2_logCHR-CD	.054*	0.038	.171**	− .337**	− .306**	.076**	.566**	.680**	1						
25. Sum_nStroke	− .069**	− .067**	− .059**	.360**	.230**	− 0.006	− .103**	− .110**	− .108**	1					
26. Sum_nRadical	− .111**	− .231**	− .175**	.356**	.465**	0.014	− .157**	− .229**	− .180**	.519**	1				
27. Sum_nPronunciation	.209**	.105**	.097**	− 0.011	0.028	.723**	.219**	.063**	.052*	.122**	0.041	1			
28. Sum_nMeaning	.586**	.365**	.370**	− .194**	− .205**	.148**	.600**	.337**	.367**	.324**	− 0.026	.337**	1		
29. Sum_ Count_Sum	.237**	.573**	.345**	− .235**	− .231**	0.024	.291**	.540**	.358**	.411**	− .077**	.191**	.707**	1	
30. Sum_logCHR_CD	.198**	.319**	.376**	− .121**	− .139**	0.012	.203**	.245**	.345**	.653**	.104**	.194**	.721**	.847**	1

RT, reaction time; ACC, accuracy rate; C1-C4, the first to forth characters of the word; nStroke, number of strokes; Sum_nStroke, nRadical, number of radicals; nPronunciation, number of pronunciations; nMeaning, number of meanings. **. Correlation is significant at the 0.01 level (two-tailed). *. Correlation is significant at the 0.05 level (two-tailed). All variables are computed after log-transformation except for the zRTs, ACC and Volume

Specifically, compared to adult participants, children who have more recent exposure to these characters and words are more sensitive to this frequency factor solely based on textbooks. Furthermore, our results showed that the factor of Count_Sum explained less variance in the naming data for both characters and words than the factor of frequency from Cai and Brybaert (2010). One possible reason is that the Count_Sum was extracted from a limited size of texts. According to the CCLOWW database which incorporates 34,671,424 character tokens and 22,427,010 word tokens (Li et al., 2022), robust frequency effect was detected in lexical processing tasks in both adults and children. Therefore, it is possible that character frequency obtained from the textbooks for elementary schools in the current study is not adequate to capture the overall distributional information.

Secondly, we found that the number of strokes had a significant effect on naming latency of characters and words in all three groups, but the number of radicals did not. As expected, characters and words with more strokes were named more slowly than those with less strokes, which replicated the stroke-number effect observed in many previous studies (for example, Just & Carpenter, 1987; Leong et al., 1987; Su & Samuels, 2010). Resembling the frequency effect, a more pronounced stroke effect was attested in the younger children group than in the other groups, showing that it was more difficult for younger children to name characters or words with more visually complex. One possibility is that these characters or words require additional cognitive resources for processing. Similarly, naming latency increased with the length of words. That is, it took longer to retrieve encoding information of words that contain two or more characters. However, it was shown in the current analysis that all three groups responded faster to words than to characters, regardless of the number of constituent characters (see Table 5 and Table 12). These findings are consistent with the U-shaped relationship between word length and response times in Tsang et al., (2018). They found that two- and three-character words induced longer RTs than single and four-character words, and that naming latency for four-character words were slightly longer than for single-character words. These findings together indicate that word processing is not necessarily more difficult than character processing, and words processing might be affected by several factors such as distributional frequency. On the other hand, the number of radicals did not show a significant effect on character naming latency in all three groups. Some studies have used both the number of strokes and radicals as indicators of the visual complexity of Chinese characters and found that both of them play significant roles in lexical processing (Liu et al., 2007; Wang & Dong, 2013; Xing et al.,

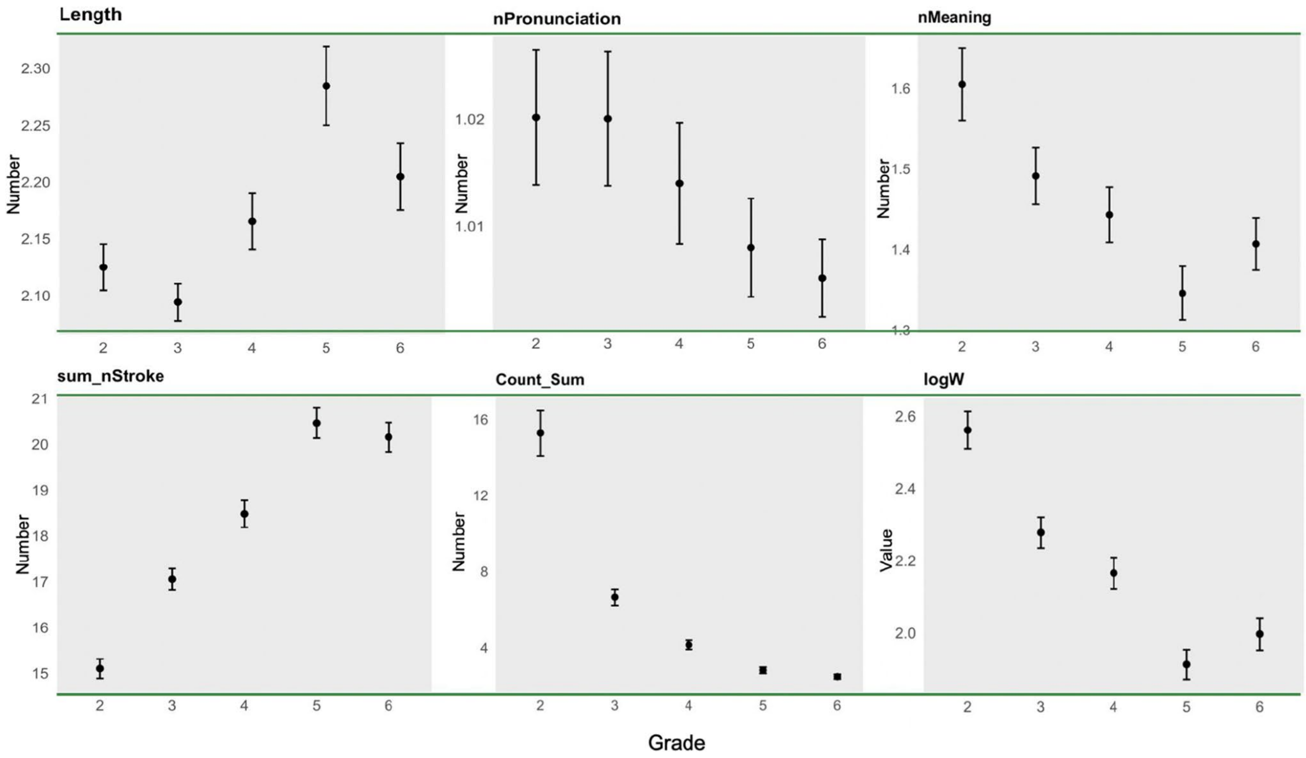

Fig. 3 Critical predictors across five grades (no first grade) in the Chinese word database. The error bar represents standard errors. Variables such as Length, nPronunciation, nMeaning, Sum_nStroke, and Count_Sum are counted as numbers; logW is log-transformed

2004). However, some other studies did not replicate this effect. For example, Su and Samuels (2010) examined the effect of the number of strokes, the number of radicals and word length on Chinese character recognition among children and adults. They ascribed the absence of the radical effect to the notion that characters with high-frequency and moderate visual complexity may not follow a radical-by-radical processing basis. In the current database, the number of strokes and radicals of characters was highly correlated ($r = .73$); therefore, it is possible that the number of strokes takes the role of the number of radicals in naming latency (Wang et al., 2020).

In addition to the effect of visual complexity on Chinese lexical processing, the structure type of Chinese characters also showed a significant effect on naming latency and accuracy in both younger and older children. In adults, however, this effect was only significant in naming latency. Specifically, naming latency for characters with top-down structure were named faster than those with left-right structure or other structures. Several studies have examined the structural effects in Chinese character processing and learning, but found inconsistent results (e.g., Li et al., 2000, 2005; Wang et al., 2020; Yu & Cao, 1992). For example, Li et al. (2000) investigated whether character recognition was affected by the character structure among school-age children (Grades 1, 3, and 5 in primary school) and adults. The structural effect was only found in Grade 1, such that characters with left-right structures showed a processing advantage relative to those with top-down or semi-enclosed structures. Tong and McBride (2014) also found that children had more errors in producing left-right structured characters and were more inclined to produce top-down structured characters. They argued that such production asymmetry is due to children's asymmetrical exposure to these two types of structures. However, in the study by Wang et al., (2020), writing latency and duration for left-right characters in adults were found to be shorter than characters with other structures, and latency and duration of top-down characters were longer than those for characters of other structures. Despite the inconsistent results, all the aforementioned studies indicate an effect of character structure on character processing. These divergent findings may be attributable to differences in tasks and reading experience among participants, and thus call for further empirical evidence.

Thirdly, as the essential role of grapheme-to-phoneme correspondence (GPC) rules in word recognition and reading acquisition observed in languages with regular and consistent orthographies (e.g. Seidenberg et al., 1984; Ziegler et al., 2003, 2010), the print-to-sound mapping at the sublexical level in Chinese is also important in the early stages of reading (Li et al., 2018; Shu et al., 2000).

Table 13 Results of stepwise MRA on zRT and ACC in the younger children group on the word level

	B	*t*	*p*	ΔR^2	*lmg*
zRT					
Intercept	0.065	1.471	0.142		
Volume	0.020	8.908	0.000	0.089	26.13
logW-CD	– 0.097	– 14.301	0.000	0.154	30.78
C1_CS	– 0.079	– 7.165	0.000	0.019	12.86
C1_nMeaning	– 0.046	– 2.021	0.043	0.001	3.39
C2_CS	– 0.061	– 6.316	0.000	0.007	9.23
C2_nPro	0.136	3.124	0.002	0.003	0.92
Sum_nStr	0.048	2.564	0.010	0.025	13.30
Sum_logCHR_CD	0.021	4.416	0.000	0.006	3.40
R^2 = .304					
ACC					
Intercept	0.747	23.319	0.000		
Volume	– 0.009	– 7.487	0.000	0.078	35.41
logW-CD	0.025	6.746	0.000	0.038	19.64
C1_logCHR-CD	0.030	3.555	0.000	0.014	12.00
C1_CS	0.020	3.039	0.002	0.002	10.76
C1_nPro	– 0.064	– 1.995	0.046	0.002	3.56
C1_nstr	0.047	2.823	0.005	0.003	1.19
C2_CS	0.021	4.624	0.000	0.011	11.83
Sum_nPro	– 0.067	– 3.381	0.001	0.011	5.61
R^2 = .159					

C1, C2, the first and second character of the word; C1_logCHR-CD, logCHR-CD of C1; C1_CS, Count_Sum of C1; C1_nstr, number of strokes of C1; C1_nMeaning, number of meanings of C1; C1_nPro, number of pronunciations of C1; C2_CS, Count_Sum of C2; C2_nPro, number of pronunciations of C2; Sum_nStr, the total strokes of the word; Sum_nPro, the total pronunciations of the component characters of the word; Sum_logCHR_CD, the total logCHR_CD of the component characters of the word. *lmg* Measures (relative importance) are scaled to 100%

Table 14 Results of stepwise MRA on zRT and ACC in the older children group on the word level

	B	*t*	*p*	ΔR^2	*lmg*
zRT					
Intercept	0.018	0.481	0.630		
Volume	0.004	2.582	0.010	0.002	9.55
logW-CD	– 0.062	– 14.013	0.000	0.184	32.19
C1_logCHR-CD	– 0.054	– 4.883	0.000	0.004	16.90
C1_CS	– 0.060	– 7.985	0.000	0.092	17.08
C1_nRa	0.041	1.981	0.048	0.001	3.04
C2_CS	– 0.036	– 5.850	0.000	0.009	6.14
Sum_nStr	0.030	2.343	0.019	0.029	11.36
Sum_nPro	0.054	3.134	0.002	0.004	0.91
Sum_logCHR_CD	0.011	3.305	0.001	0.002	2.82
R^2 = .328					
ACC					
Volume	– 0.002	– 2.806	0.005	0.004	16.73
logW-CD	0.012	6.160	0.000	0.031	23.69
C1_logCHR-CD	0.010	2.660	0.008	0.003	11.99
C1_nPro	– 0.038	– 2.118	0.034	0.002	12.83
Sum_CS	0.005	4.419	0.000	0.022	11.44
Sum_nPro	– 0.059	– 5.310	0.000	0.029	23.31
R^2 = .090					

C1, C2, the first and second character of the word; C1_logCHR-CD, logCHR-CD of C1; C1_CS, Count_Sum of C1; C1_nRa, number of radicals of C1; C1_nPro, number of pronunciations of C1; C2_CS, Count_Sum of C2; Sum_nStr: the total strokes of the word; Sum_nPro: the total pronunciations of the component characters of the word; Sum_CS: the total Count_Sum of the component characters of the word; Sum_logCHR_CD: the total logCHR_CD of the component characters of the word. lmg Measures (relative importance) are scaled to 100%

The result of the current study showed that characters with regular pronunciation were named faster than those irregular, or semi-regular phonograms in both younger and older children. As for adults, characters with semi-regular pronunciation (including semi-regular, rhyming, and alliteration), were named more slowly relative to regular and irregular phonograms. This was consistent with previous findings (Chang et al., 2016; Liu et al., 2007; Wang et al., 2020), and further supported that phonological activation of phonetic radicals in phonograms would interfere character naming. Shu and Meng (1996) found that children performed significantly better for regular characters than for irregular characters and those with unpronounceable radicals. They claimed that children were able to make use of phonological cues of character radicals in naming tasks, especially when these characters were unfamiliar to them. Liu et al. (2006) also suggested that lower graders (Grades 1 and 2) with an earlier awareness of character regularity were able to develop better literacy. Note that the knowledge of how phonological cues can be used to access character pronunciations is not explicitly taught in school. Children typically acquire orthographic and phonological knowledge about characters based on reading experience. In addition, we found that approximately 31% of all characters in the current database were phonograms with reliable phonological cues (lexical tones are not considered) provided by phonetic radicals, compared to 39% in Shu et al. (2003). Adults are more likely to reduce their reliance on phonological clues provided by phonetic radicals when naming phonograms as their reading experience increases. This might also explain the reason why the advantage of naming regular characters was absent in adults.

Last but not least, we found that both the number of pronunciations and meanings of characters and words had significant effects on lexical processing, but in the opposite way. Specifically, consistent with Tsang et al. (2018), characters with multiple pronunciations were named more

Table 15 Results of stepwise MRA on zRT and ACC in the adult group on the word level

	B	*t*	*p*	*ΔR*^{2}	*lmg*
zRT					
Intercept	– 0.014	– 0.522	0.602		
Volume	0.003	2.448	0.014	0.007	8.8
Length	0.020	2.087	0.037	0.002	8.4
logW-CD	– 0.034	– 10.012	0.000	0.128	33.4
C1_logCHR-CD	– 0.041	– 5.046	0.000	0.038	19.6
C1_CS	– 0.013	– 2.229	0.026	0.002	9.0
C1_nMeaning	– 0.022	– 2.010	0.045	0.002	5.8
C2_CS	– 0.013	– 3.219	0.001	0.003	4.2
Sum_nStr	0.020	2.052	0.040	0.024	9.8
Sum_nPro	0.032	2.355	0.019	0.002	1.1
R^2 = .206					
ACC					
Intercept	0.959	170.949	0.000		
logW-CD	0.003	3.688	0.000	0.015	21.4
C1_CS	0.003	2.847	0.004	0.004	7.2
C2_logCHR-CD	0.006	3.861	0.000	0.007	19.0
Sum_nPro	– 0.030	– 7.753	0.000	0.025	52.4
R^2 = .050					

C1, C2, the first and second character of the word; C1_logCHR-CD, logCHR-CD of C1; C1_CS, Count_Sum of C1; C1_nMeaning, number of meanings of C1; C2_logCHR-CD, logCHR-CD of C2; C2_CS, Count_Sum of C2; Sum_nStr: the total strokes of the word; Sum_nPro: the total pronunciations of the component characters of the word. *lmg* Measures (relative importance) are scaled to 100%

slowly and less accurately, indicating a phonological interference in character naming (Tan & Peng, 1990). In addition, this effect was more pronounced in adults than in older children but was absent in younger children. One possible reason is that different pronunciations (often with different meanings) of the same character are often distributed in different grades, and younger children who have not been exposed to all pronunciations or meanings are less likely to be affected by the interference of character pronunciations in the naming task. By contrast, characters with multiple meanings and words with semantically rich characters (especially when those characters appear in the initial position) were named much faster. Many previous studies have also demonstrated that characters with more than one meaning were processed faster than semantically unambiguous characters, known as 'ambiguity advantage' (Lee et al., 2015; Chang et al., 2016; Lin & Ahrens, 2010). Note that the ambiguity advantage in our analysis of naming latency was only found in older children and adults. For naming accuracy, this effect was only significant for younger children. A possible reason is that younger children have not yet acquired multiple meanings for some common characters. On the other hand, significant ambiguity advantage in naming accuracy for the younger children group might be attributed to other factors such as frequency (Count_Sum and logCHR-CD), given that multi-meaning characters are more likely to have high frequency (*r* = .465 for Count_Sum and .522 for logCHR-CD, see Table 5).

Taken together, the current results indicate that different lexical variables may have different effects for children and adults. Specifically, the orthography-related factors, such as visual complexity and structural types of characters, play more significant roles for children who are at the early stages of learning. The phonological and semantic effects in character processing might come into play as literacy and reading experience increase.

Table 16 Results of regression coefficient comparisons on zRT among three groups on the word level

Variables	Group 1	Group 2	b1	b2	difference	*t*	*p*
Volume	Younger children	Older children	0.021	0.004	0.017	7.268	< 0.001
	Younger children	Adult	0.021	0.003	0.018	– 3.483	< 0.001
	Older children	Adult	0.004	0.003	0.001	0.775	0.438
logW-CD	Younger children	Older children	– 0.095	– 0.069	– 0.026	– 1.847	0.065
	Younger children	Adult	– 0.095	– 0.045	– 0.049	3.166	0.002
	Older children	Adult	– 0.069	– 0.045	– 0.023	8.038	< 0.001
C1_CS	Younger children	Older children	– 0.069	– 0.077	0.008	– 6.979	< 0.001
	Younger children	Adult	– 0.069	– 0.034	– 0.035	– 3.791	< 0.001
	Older children	Adult	– 0.077	– 0.034	– 0.043	– 3.517	< 0.001
C2_CS	Younger children	Older children	– 0.039	– 0.022	– 0.018	5.503	< 0.001
	Younger children	Adult	– 0.039	– 0.008	– 0.032	0.53	0.596
	Older children	Adult	– 0.022	– 0.008	– 0.014	– 4.759	< 0.001
Sum_nStr	Younger children	Older children	0.108	0.064	0.045	– 6.645	< 0.001
	Younger children	Adult	0.108	0.035	0.073	– 2.259	0.024
	Older children	Adult	0.064	0.035	0.029	3.113	0.002

Table 17 Results of regression coefficient comparisons on ACC among three groups on the word level

Variables	Group 1	Group 2	b1	b2	difference	*t*	*p*
logW-CD	Younger children	Older children	0.036	0.014	0.022	5.716	< 0.001
	Younger children	Adult	0.036	0.007	0.029	4.278	< 0.002
	Older children	Adult	0.014	0.007	0.007	8.04	< 0.003
C2_CS	Younger children	Older children	0.032	0.011	0.021	6.612	< 0.004
	Younger children	Adult	0.032	0.002	0.03	3.043	0.002
	Older children	Adult	0.011	0.002	0.009	3.27	0.001

Age of acquisition

The current database also extended a number of studies concerned with the age of acquisition of Chinese lexical units (Cai et al., 2021; Chang et al., 2016; Chang & Lee, 2020; Sze et al., 2014; Wang et al., 2020; Xu et al., 2020) by using objective measures for characters and words based on textbooks used in elementary schools. The volume information was used to represent the time when these characters and words were typically learned, same as Cai et al (2021) and Liu et al. (2007). The current results showed that the factor Volume significantly affected character and word naming across all three groups. Moreover, the contribution of Volume in character naming performance was second only to frequency in all three groups. These results were consistent with existing evidence across different languages, showing that AoA plays a crucial role in lexical decision (Bylund et al., 2019; Chang & Lee, 2020; Chen et al., 2009a, b), word-naming (Bonin et al., 2001; Bylund et al., 2019; Chang & Lee, 2020; Chen et al., 2004; Liu et al., 2007), and word-writing tasks (Bonin et al., 2001; Wang et al., 2020), indicating that early acquired characters and words can also be recognized, produced or written more quickly and accurately (Brysbaert & Cortese, 2011; Chang & Lee, 2020; Ferrand, 2011; Sze et al., 2014).

Moreover, AoA effect was consistently observed in character and word naming in both adults and children. However, the AoA effect in lexical processing demonstrated different patterns among the three groups. The overall pattern was that the effect decreased with age; we found that the AoA effect on word naming was less significant in older children and adults than in younger children. Cumulative-frequency hypothesis postulates that the difference of cumulative frequency between early-acquired words and late-acquired words narrows with age, so younger children are more sensitive to the AoA effect (Ghyselinck et al., 2004; Lewis et al., 2001). Chen et al. (2004) also found that both AoA and frequency had significant effects on naming Chinese disyllabic words; they claimed that AoA was related to semantic processing because stronger and more reliable AoA effect was observed in semantic-related tasks, such as word association and categorization (Brysbaert et al., 2000; Van Loon-Vervoorn & Willemsen, 1989). Therefore, we speculate that word processing may also be affected by the interaction between AoA and semantics. Furthermore, written AoA, rather than spoken AoA (Cai et al., 2021; Xu et al., 2020), was investigated in the current study. For Chinese children, individual characters are not commonly used in spoken language and usually acquired by classroom-learning at school age (McBride-Chang & Ho, 2000), while they are more likely to have acquired words' spoken forms prior to the school education. Therefore, the spoken AoA might be more influential for word processing. From this perspective, it is reasonable why the current written AoA had different effects on naming characters and words.

Conclusion

In conclusion, the present study extends the current body of studies by establishing a database of Chinese characters and words based on the newly issued elementary school textbooks used in mainland China. For all characters and words to be learned in elementary schools, we extracted key lexical variables from either the textbooks or external resources and obtained naming latency and accuracy from children and adults to measure character and word processing. Our database offers a nuanced view of the development of lexical variables as grade increases and provides an empirical basis for a better understanding of the effects of those variables in lexical processing among school-age children and adults. This attempt will facilitate studies in Chinese language acquisition by showing normative data and distributional information of lexical units in terms of orthography, phonology, semantics, and other aspects.

Acknowledgments The first two authors contributed equally to the manuscript. We thank Leshan Chen for inspiring the research idea. The database is available on Open Science Framework (OSF) (https://osf.io/ynk8c/?view_only=5186bd68549340bd923e9b6531d2c820).

Funding The study was supported by the National Natural Science Foundation of China (31871097), the National Key Basic Research Program of China (2014CB846102), the Interdisciplinary Research Funds of Beijing Normal University, and the Fundamental Research Funds for the Central Universities (2017XTCX04).

Declarations

Conflict of interest None declared.

Reference

Arciuli, J., & Simpson, I. C. (2012). Statistical learning is related to reading ability in children and adults. *Cognitive Science, 36*(2), 286–304. https://doi.org/10.1111/j.1551-6709.2011.01200.x

Baron, J., & Strawson, C. (1976). Use of orthographic and word-specific knowledge in reading words aloud. *Journal of Experimental Psychology: Human Perception and Performance, 2*(3), 386–393.

Beijing Academy of Educational Sciences. (1998). *Liunianzhi xiaoxue shiyong keben* (Elementary school textbooks for first through sixth grades). Beijing Press.

Bonin, P., Fayol, M., & Chalard, M. (2001). Age of acquisition and word frequency in written picture naming. *The Quarterly Journal of Experimental Psychology Section A, 54*(2), 469–489. https://doi.org/10.1080/713755968

Borowsky, R., & Masson, M. E. J. (1996). Semantic ambiguity effects in word identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(1), 63–85.

Brady, S. A., & Shankweiler, D. (1991). *Phonological processes in literacy: a tribute to Isabelle Y. Liberman*. Erlbaum Associates.

Brysbaert, M., & Cortese, M. J. (2011). Do the effects of subjective frequency and age of acquisition survive better word frequency norms? *Quarterly Journal of Experimental Psychology, 64*(3), 545–559. https://doi.org/10.1080/17470218.2010.503374

Brysbaert, M., Wijnendaele, I. V., & Deyne, S. D. (2000). Age-of-acquisition effects in semantic processing tasks. *Acta Psychologica, 104*, 215–226.

Bylund, E., Abrahamsson, N., Hyltenstam, K., & Norrman, G. (2019). Revisiting the bilingual lexical deficit: The impact of age of acquisition. *Cognition, 182*, 45–49. https://doi.org/10.1016/j.cognition.2018.08.020

Byrne, B. (1992). Studies in the acquisition procedure for reading: Rationale, hypotheses, and data. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 1–34). Erlbaum.

Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PLoS ONE, 5*(6), e10729. https://doi.org/10.1371/journal.pone.0010729

Cai, Z. G., Huang, S., Xu, Z., & Zhao, N. (2021). Objective ages of acquisition for 3300+ simplified Chinese characters. *Behavior Research Methods, 54*(1), 311–323. https://doi.org/10.3758/s13428-021-01626-1

Carroll, J. B., & White, M. N. (1973). Word frequency and age of acquisition as determiners of picture-naming latency. *Quarterly Journal of Experimental Psychology, 25*(1), 85–95. https://doi.org/10.1080/14640747308400325

Chang, Y.-N., & Lee, C.-Y. (2020). Age of acquisition effects on traditional Chinese character naming and lexical decision. *Psychonomic Bulletin & Review, 27*(6), 1317–1324. https://doi.org/10.3758/s13423-020-01787-8

Chang, Y.-N., Hsu, C.-H., Tsai, J.-L., Chen, C.-L., & Lee, C.-Y. (2016). A psycholinguistic database for traditional Chinese character naming. *Behavior Research Methods, 48*(1), 112–122. https://doi.org/10.3758/s13428-014-0559-7

Chen, Y. P., Allport, D. A., & Marshall, J. C. (1996). What are the functional orthographic units in chinese word recognition: The stroke or the stroke pattern? *The Quarterly Journal of Experimental Psychology Section A, 49*(4), 1024–1043. https://doi.org/10.1080/713755668

Chen, B., Wang, L., Wang, L., & Peng, D. (2004). The effect of age of word acquisition and frequency on the identification of Chinese double-character words. *Psychological Science, 27*(5), 1060–1064.

Chen, M.-J., Weekes, B. S., Peng, D., & Lei, Q. (2006). Effects of semantic radical consistency and combinability on Chinese character processing. In P. Li, L. H. Tan, E. Bates, & O. J. L. Tzeng (Eds.), *The Handbook of East Asian Psycholinguistics* (Vol. 1, pp. 175–186). Cambridge University Press.

Chen, B., Dent, K., You, W., & Wu, G. (2009). Age of acquisition affects early orthographic processing during Chinese character recognition. *Acta Psychologica Sinica, 130*(3), 196–203. https://doi.org/10.1016/j.actpsy.2008.12.004

Chen, X., Hao, M., Geva, E., Zhu, J., & Shu, H. (2009). The role of compound awareness in Chinese children's vocabulary acquisition and character reading. *Reading and Writing, 22*(5), 615–631. https://doi.org/10.1007/s11145-008-9127-9

Chinese Academy of Social Sciences. (2012). *Xiandai Hanyu Cidian* (现代汉语词典). Commercial Press.

Corral, S., Ferrero, M., & Goikoetxea, E. (2009). LEXIN: A lexical database from Spanish kindergarten and first-grade readers. *Behavior Research Methods, 41*(4), 1009–1017. https://doi.org/10.3758/BRM.41.4.1009

De Francis, J. (1989). *Visible speech: The diverse oneness of writing systems*. University of Hawaii Press.

Elementary Education Teaching and Research Center, Beijing Education and Science Institute. (1996). *Elementary school textbooks.* Beijing, China: Beijing Publishers.

Fan, K. Y., Gao, J. Y., & Ao, X. P. (1984). Pronunciation principles of the Chinese character and alphabetic writing scripts. *Chinese Character Reform, 3*, 23–27.

Fang, S.-P. (1994). English word length effects and the Chinese character-word difference: Truth or myth? *Chinese Journal of Psychology, 36*(1), 59–79.

Feldman, L. B., & Siok, W. W. T. (1999). Semantic radicals contribute to the visual identification of chinese characters. *Journal of Memory and Language, 40*(4), 559–576. https://doi.org/10.1006/jmla.1998.2629

Ferrand, L. (2011). Comparing word processing times in naming, lexical decision, and progressive demasking: Evidence from Chronolex. *Frontiers in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00306

Fu, Y. (1989). *Dictionary of Chinese character properties* (汉字属性字典). Language and Culture Press.

Ghyselinck, M., Lewis, M. B., & Brysbaert, M. (2004). Age of acquisition and the cumulative-frequency hypothesis: A review of the literature and a new multi-task investigation. *Acta Psychologica, 115*(1), 43–67. https://doi.org/10.1016/j.actpsy.2003.11.002

Gilhooly, K. J., & Logie, R. H. (1980). Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. *Behavior Research Methods & Instrumentation, 12*(4), 395–427. https://doi.org/10.3758/BF03201693

Goswami, U., & Bryant, P. E. (1990). *Phonological skills and learning to read*. Erlbaum.

Grömping U (2006). *Relaimpo: Relative Importance of Regressors in Linear Models*. R package version 1.1-1.

Hair, J. F. (2011). Multivariate Data Analysis: An Overview. In M. Lovric (Ed.), *International Encyclopedia of Statistical Science.* Springer. https://doi.org/10.1007/978-3-642-04898-2_395

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

He, K., & Li, D. (1987). *Xiandai Hanyu san qian changyong ci biao* [Three thousand most commonly used words in modern Chinese]. Beijing Normal University Press.

He, X., Xue, J., & Shu, H. (2011). The effect of regularity and transparency on Chinese characters output by those with dyslexia:

From the perspective of connectionists' reading model. *Chinese Journal of Special Education, 6*, 37–41.

Ho, C.S.-H., Ng, T.-T., & Ng, W.-K. (2003). A "radical" approach to reading development in Chinese: The role of semantic radicals and phonetic radicals. *Journal of Literacy Research, 35*(3), 849–878. https://doi.org/10.1207/s15548430jlr3503_3

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), 1–11. https://doi.org/10.1016/j.bandl.2010.10.002

Huang, C. R., & Chen, K. J. (1998). *Academia Sinica balanced corpus* (version 3). Taipei, Taiwan: Academia Sinica.

Huang, X., Lin, D., Yang, Y., Xu, Y., Chen, Q., & Tanenhaus, M. K. (2021). Effects of character and word contextual diversity in Chinese beginning readers. *Scientific Studies of Reading, 25*(3), 251–271. https://doi.org/10.1080/10888438.2020.1768258

Ismail, N., & Jemain, A. A. (2007). Handling overdispersion with negative binomial and generalized Poisson regression models. *Casualty actuarial society forum* (2007th ed., pp. 103–58). Citeseer.

Johnston, R. A., & Barry, C. (2006). Age of acquisition and lexical processing. *Visual Cognition, 13*(7–8), 789–845. https://doi.org/10.1080/13506280544000066

Just, M. A., & Carpenter, P. A. (1987). *The psychology of reading and language comprehension.* Allyn & Bacon.

Kaschak, M. P. (2007). Long-term structural priming affects subsequent patterns of language production. *Memory & Cognition, 35*(5), 925–937. https://doi.org/10.3758/BF03193466

Katz, L., Brancazio, L., Irwin, J., Katz, S., Magnuson, J., & Whalen, D. H. (2012). What lexical decision and naming tell us about reading. *Reading and Writing, 25*(6), 1259–1282. https://doi.org/10.1007/s11145-011-9316-9

Kellas, G., Ferraro, F. R., & Simpson, G. B. (1988). Lexical ambiguity and the timecourse of attentional allocation in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 14*(4), 601–609.

Lambert, É., & Chesnet, D. (2001). NOVLEX: Une base de données lexicales pour les élèves de primaire. *L'année psychologique, 101*(2), 277–288. https://doi.org/10.3406/psy.2001.29557

Lee, C.-Y., Hsu, C.-H., Chang, Y.-N., Chen, W.-F., & Chao, P.-C. (2015). the feedback consistency effect in chinese character recognition: Evidence from a psycholinguistic norm. *Language and Linguistics, 16*(4), 535–554. https://doi.org/10.1177/1606822X15583238

Lei, L., Pan, J., Liu, H., McBride-Chang, C., Li, H., Zhang, Y., et al. (2011). Developmental trajectories of reading development and impairment from ages 3 to 8 years in Chinese children. *Journal of Child Psychology and Psychiatry, 52*(2), 212–220. https://doi.org/10.1111/j.1469-7610.2010.02311.x

Leong, C. K., Cheng, P.-W., & Mulcahy, R. (1987). Automatic processing of morphemic orthography by mature readers. *Language and Speech, 30*(2), 181–196. https://doi.org/10.1177/002383098703000207

Lété, B., Sprenger-Charolles, L., & Colé, P. (2004). MANULEX: A grade-level lexical database from French elementary school readers. *Behavior Research Methods, Instruments, & Computers, 36*(1), 156–166. https://doi.org/10.3758/BF03195560

Lewis, M. B., Gerhand, S., & Ellis, H. D. (2001). Re-evaluating age-of-acquisition effects: Are they simply cumulative-frequency effects? *Cognition, 78*(2), 189–205. https://doi.org/10.1016/S0010-0277(00)00117-7

Lexicographical Center of Commercial Press. (2002). *Xinhua idiom dictionary* (新华成语词典*)*. Commercial Press.

Li, D. (1993). *A study of Chinese characters*. Peking University Press.

Li, J., Fu, X., & Lin, Z. (2000). Study on the development of Chinese orthographic regularity in school children. *Acta Psychologica Sinica, 32*(2), 121–126.

Li, L. H., Liu, H. G., & Liu, X. L. (2005). Effects of characters construction on basic processing unit of Chinese character recognition. *Psychological exploration, 25*, 23–27.

Li, H., Shu, H., McBride-Chang, C., Liu, H., & Peng, H. (2012). Chinese children's character recognition: Visuo-orthographic, phonological processing and morphological skills: CHINESE CHILDREN'S CHARACTER RECOGNITION. *Journal of Research in Reading, 35*(3), 287–307. https://doi.org/10.1111/j.1467-9817.2010.01460.x

Li, L., Wang, H. C., Castles, A., Hsieh, M. L., & Marinus, E. (2018). Phonetic radicals, not phonological coding systems, support orthographic learning via self-teaching in Chinese. *Cognition, 176*, 184–194.

Li, M.-F., Gao, X.-Y., & Wu, J.-T. (2020). Neighborhood effects in Chinese character recognition: Going beyond phonological perspectives to explain a possible underlying mechanism. *Reading and Writing, 33*(3), 547–570. https://doi.org/10.1007/s11145-019-09973-4

Li, L., Yang, Y., Song, M., Fang, S., Zhang, M., Chen, Q., Cai, Q. (2022). CCLOWW: A grade-level Chinese children's lexicon of written words. *Behavior Research Methods*, 1–16. https://doi.org/10.3758/s13428-022-01890-9

Li, H. (2018). An analysis of the characteristics and learning applicability of the nationally compiled Chinese textbooks for primary schools (国家统编小学语文教科书的特色与学习适用性分析). *Educational Science Research*, *8* (In Chinese)

Lin, Chien-Jer Charles., & Ahrens, Kathleen. (2010). Ambiguity advantage revisited: Two meanings are better than one when accessing Chinese nouns. *Journal of Psycholinguistic Research, 39*, 1–19. https://doi.org/10.1007/s10936-009-9120-8

Linguistics Institute of Chinese Academy of Social Sciences. (2020). *Xinhua dictionary (*新华字典*, version 12)*. Commercial Press.

Liu, Y., Shu, H., & Li, P. (2007). Word naming and psycholinguistic norms: Chinese. *Behavior Research Methods, 39*(2), 192–198.

Liu, X., Liu, W., Zhang, L., Xu, X., Zhang, W., Zhang, X., Zhang, J. (2006). A study on the relationship between children's literacy and character regularity awareness. *Chinese Journal of Special Education, 61*(7), 56–61.

Lu, S. C. (1989). Frequency dictionary of Chinese characters, words and phrases used in Singapore primary school textbooks. *Chinese Language and Research Centre, National University of Singapore, Singapore.*

Lu, S. C. (1992). Frequency dictionary of Chinese characters, words and phrases used in Singapore secondary school textbooks. *Chinese Language and Research Centre, National University of Singapore, Singapore.*

Luo, Z. (1986). *The Great Chinese Word Dictionary*(汉语大词典, version 1). Shanghai Lexicographical Publishing House.

Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J. R., Bethard, S., McClosky, D. (2014). *The Stanford CoreNLP natural language processing toolkit*. Paper presented at the 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations, Baltimore, Maryland, USA.

Marconi, L., Ott, M., Pesenti, E., Ratti, D., & Tavella, M. (1993). *Lessico elementare: Dati statistici sull'italiano scritto e letto dai bambini delle elementari [Elementary lexicon: Statistical data for Italian written and spoken by elementary school children]*. Zanichelli.

Martín, J. A. M., & Pérez, M. E. G. (2008). ONESC: A database of orthographic neighbors for Spanish read by children. *Behavior Research Methods, 40*(1), 191–197. https://doi.org/10.3758/BRM.40.1.191

Masterson, J., Stuart, M., Dixon, M., & Lovejoy, S. (2010). Children's printed word database: Continuities and changes over time in children's early reading vocabulary. *British Journal of Psychology, 101*(2), 221–242. https://doi.org/10.1348/000712608X371744

McBride, C. A. (2016). Is Chinese special? Four aspects of chinese literacy acquisition that might distinguish learning Chinese from learning alphabetic orthographies. *Educational Psychology Review, 28*(3), 523–549. https://doi.org/10.1007/s10648-015-9318-2

McBride-Chang, C., & Ho, C. S. H. (2000). Developmental issues in Chinese children's character acquisition. *Journal of Educational Psychology, 92*(1), 50–55. https://doi.org/10.1037/0022-0663.92.1.50

Meng, X., Shu, H., & Zhou, X. (2000). Children's Chinese character structure awareness in character output. *Psychological Science, 23*(3), 260–240.

Meng, X., Shu, H., Zhou, X., & Luo, X. (2000). Character production and recognition in Chinese processing: a comparative study between poor readers and normal readers of fourth grade. *Acta Psychologica Sinica, 32*(02), 133–138.

Morrison, C. M., & Ellis, A. W. (2000). Real age of acquisition effects in word naming and lexical decision. *British Journal of Psychology, 91*(2), 167–180. https://doi.org/10.1348/000712600161763

Morrison, C. M., Chappell, T. D., & Ellis, A. W. (1997). Age of acquisition norms for a large set of object names and their relation to adult estimates and other variables. *The Quarterly Journal of Experimental Psychology Section A, 50*(3), 528–559. https://doi.org/10.1080/027249897392017

Myers, J. (2019). *The grammar of Chinese characters: Productive knowledge of formal patterns in an orthographic system*. Routledge.

National Language Commission, Ministry of Education, PRC. (2011–2018). *Language Situation in China*. Beijing: The Commercial Press.

O'brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity, 41*(5), 673–690. https://doi.org/10.1007/s11135-006-9018-6

Packard, J. L., Chen, X., Li, W., Wu, X., Gaffney, J. S., Li, H., & Anderson, R. C. (2006). Explicit instruction in orthographic structure and word morphology helps Chinese children learn to write characters. *Reading and Writing, 19*(5), 457–487. https://doi.org/10.1007/s11145-006-9003-4

Parkin, A. J. (1982). Phonological recoding in lexical decision: Effects of spelling-to-sound regularity depend on how regularity is defined. *Memory & Cognition, 10*(1), 43–53. https://doi.org/10.3758/BF03197624

Peng, D., & Wang, C. (1997). Basic processing unit of Chinese character recognition: Evidence from stroke number effect and radical number effect. *Acta Psychologica Sinica Acta Psychologica Sinica, 29*(1), 8–16.

Peng, D., Deng, Y., & Chen, B. (2003). The polysemy effect in Chinese one-character word identification. *Acta Psychologica Sinica, 35*(5), 569–575.

Pérez, M. A. (2007). Age of acquisition persists as the main factor in picture naming when cumulative word frequency and frequency trajectory are controlled. *Quarterly Journal of Experimental Psychology, 60*(1), 32–42. https://doi.org/10.1080/17470210600577423

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*(1), 101–118.

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences, 10*(5), 233–238. https://doi.org/10.1016/j.tics.2006.03.006

Plonsky, L., & Ghanbar, H. (2018). Multiple regression in L2 research: A methodological synthesis and guide to interpreting R2 values. *The Modern Language Journal, 102*(4), 713–731.

Que, D. L. (2008). *Longman Chinese dictionary*. Hong Kong: Longman.

Raven, J. C., Court, J. H., & Raven, J. (1983). *Manual for Ravens Progressive Matrices and Vocabulary Scales*. Section 3: Standard Progressive Matrices.

Rodd, J., Gaskell, G., & Marslen-Wilson, W. (2002). Making Sense of Semantic Ambiguity: Semantic Competition in Lexical Access. *Journal of Memory and Language, 46*(2), 245–266. https://doi.org/10.1006/jmla.2001.2810

Romberg, A. R., & Saffran, J. R. (2010). Statistical learning and language acquisition. *WIREs Cognitive Science, 1*(6), 906–914. https://doi.org/10.1002/wcs.78

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*(5294), 1926–1928. https://doi.org/10.1126/science.274.5294.1926

Schroeder, S., Würzner, K.-M., Heister, J., Geyken, A., & Kliegl, R. (2015). childLex: A lexical database of German read by children. *Behavior Research Methods, 47*(4), 1085–1094. https://doi.org/10.3758/s13428-014-0528-1

Seidenberg, M. S., Waters, G. S., Barnes, M. A., & Tanenhaus, M. K. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behavior, 23*(3), 383–404. https://doi.org/10.1016/S0022-5371(84)90270-6

Shu, H., & Meng, X. (1996). Awareness of phonological cues in pronunciation of Chinese characters and its development. *Acta Psychologica Sinica, 28*(2), 160–165.

Shu, H., Anderson, R. C., & Wu, N. (2000). Phonetic awareness: Knowledge of orthography-phonology relationships in the character acquisition of Chinese children. *Journal of Educational Psychology, 92*(1), 56–62. https://doi.org/10.1037/0022-0663.92.1.56

Shu, H., Chen, X., Anderson, R. C., Wu, N., & Xuan, Y. (2003). Properties of school Chinese: Implications for learning to read. *Child Development, 74*(1), 27–47.

Soares, A. P., Medeiros, J. C., Simões, A., Machado, J., Costa, A., Iriarte, Á., de Almeida, J. J., Pinheiro, A. P., & Comesaña, M. (2014). ESCOLEX: A grade-level lexical database from European Portuguese elementary to middle school textbooks. *Behavior Research Methods, 46*(1), 240–253. https://doi.org/10.3758/s13428-013-0350-1

Song, S., Su, M., Kang, C., Liu, H., Zhang, Y., McBride-Chang, C., Tardif, T., Li, H., Liang, W., Zhang, Z., & Shu, H. (2015). Tracing children's vocabulary development from preschool through the school-age years: An 8-year longitudinal study. *Developmental Science, 18*(1), 119–131. https://doi.org/10.1111/desc.12190

Spencer, M., Kaschak, M. P., Jones, J. L., & Lonigan, C. J. (2015). Statistical learning is related to early literacy-related skills. *Reading and Writing, 28*(4), 467–490. https://doi.org/10.1007/s11145-014-9533-0

Stadthagen-Gonzalez, H., & Davis, C. J. (2006). The Bristol norms for age of acquisition, imageability, and familiarity. *Behavior Research Methods, 38*(4), 598–605. https://doi.org/10.3758/BF03193891

Stuart, M., Dixon, M., Masterson, J., & Gray, B. (2003). Children's early reading vocabulary: Description and word frequency lists. *British Journal of Educational Psychology, 73*(4), 585–598. https://doi.org/10.1348/000709903322591253

Su, Y.-F., & Samuels, S. J. (2010). Developmental changes in character-complexity and word-length effects when reading Chinese script. *Reading and Writing, 23*(9), 1085–1108. https://doi.org/10.1007/s11145-009-9197-3

Sun, H. L., Huang, J. P., Sun, D. J., Li, D. J., & Xing, H. B. (1997). Introduction to language corpus system of modern Chinese study. In *Paper collection for the fifth world Chinese teaching symposium* (pp. 459–466). Beijing: Peking University Press.

Sun, C. C., Hendrix, P., Ma, J., & Baayen, R. H. (2018). Chinese lexical database (CLD): A large-scale lexical database for simplified

Mandarin Chinese. *Behavior Research Methods, 50*(6), 2606–2629. https://doi.org/10.3758/s13428-018-1038-3

Sze, W. P., Rickard Liow, S. J., & Yap, M. J. (2014). The Chinese Lexicon Project: A repository of lexical decision behavioral responses for 2,500 Chinese characters. *Behavior Research Methods, 46*(1), 263–273. https://doi.org/10.3758/s13428-013-0355-9

Tan, L. H., & Peng, D.-L. (1990). The effects of semantic context on the feature analyses of single Chinese characters. *Acta Psychologica Sinica, 4*, 5–10.

Tan, L. H., & Perfetti, C. A. (1997). Visual Chinese Character Recognition: Does Phonological Information Mediate Access to Meaning? *Journal of Memory and Language, 37*(1), 41–57. https://doi.org/10.1006/jmla.1997.2508

Tan, L. H., Hoosain, R., & Peng, D. (1995). Role of early presemantic phonological code in Chinese character identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(1), 43–54.

Tong, X., & McBride, C. (2014). Chinese children's statistical learning of orthographic regularities: Positional constraints and character structure. *Scientific Studies of Reading, 18*(4), 291–308. https://doi.org/10.1080/10888438.2014.884098

Tong, X., McBride-Chang, C., Shu, H., & Wong, A.M.-Y. (2009). morphological awareness, orthographic knowledge, and spelling errors: Keys to understanding early Chinese literacy acquisition. *Scientific Studies of Reading, 13*(5), 426–452. https://doi.org/10.1080/10888430903162910

Tsang, Y.-K., Huang, J., Lui, M., Xue, M., Chan, Y.-W.F., Wang, S., & Chen, H.-C. (2018). MELD-SCH: A megastudy of lexical decision in simplified Chinese. *Behavior Research Methods, 50*(5), 1763–1777. https://doi.org/10.3758/s13428-017-0944-0

Tse, C.-S., Yap, M. J., Chan, Y.-L., Sze, W. P., Shaoul, C., & Lin, D. (2017). The Chinese Lexicon Project: A megastudy of lexical decision performance for 25,000+ traditional Chinese two-character compound words. *Behavior Research Methods, 49*(4), 1503–1519. https://doi.org/10.3758/s13428-016-0810-5

Tzeng, O. J. L., Zhong, H. L., Hung, D. L., & Lee, W. L. (1995). Learning to be a conspirator: A tale of becoming a good Chinese reader. In B. de Gelder & J. Morais (Eds.), *Speech and reading: A comparative approach* (pp. 227–246). Erlbaum.

Van Esch, D. (2012). *Leiden weibo corpus*. Downloaded from http://lwc.daanvanesch.nl

van Loon-Vervoorn, W. A., & Willemsen, I. (1989). Selective disturbance in lexical knowledge in the elderly with or without dementia. *Tijdschrift voor gerontologie en geriatrie, 20*(2), 59–65.

Wang, Q., & Dong, Y. (2013). The N2- and N400-like effects of radicals on complex Chinese characters. *Neuroscience Letters, 548*, 301–305. https://doi.org/10.1016/j.neulet.2013.05.074

Wang, R., Huang, S., Zhou, Y., & Cai, Z. G. (2020). Chinese character handwriting: A large-scale behavioral study and a database. *Behavior Research Methods, 52*(1), 82–96. https://doi.org/10.3758/s13428-019-01206-4

Wang, H., & Chen, Q. (2019). The process, achievements and experience of the construction of Chinese language textbooks for primary school over the past 70 years (小学语文教材建设70年: 历程、成就、经验). Curriculum, Teaching Material and Method, *11*. (In Chinese)

Wen, R. (2016). The concept, characteristics and use suggestions of the Chinese teaching material compiled by Ministry of Education ("部编本"语文教材的编写理念、特色与使用建议). Curriculum, Teaching Material and Method, *11*. (In Chinese)

Wen, R. (2017). How to use the primary Chinese teaching material compiled by Ministry of Education (如何用好"部编本"小学语文教材). *Primary Chinese*, 25–31. (In Chinese)

Xing, H., Shu, H., & Li, P. (2004). The acquisition of Chinese characters: Corpus analyses and connectionist simulations. *Journal of Cognitive Science, 5*, 1–49.

Xu, X., Li, J., & Guo, S. (2020). Age of acquisition ratings for 19,716 simplified Chinese words. *Behavior Research Methods, 53*(2), 558–573. https://doi.org/10.3758/s13428-020-01455-8

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*(2), 238–257. https://doi.org/10.1016/j.jml.2009.05.001

Yang, X., Peng, P., & Meng, X. (2019). How do metalinguistic awareness, working memory, reasoning, and inhibition contribute to Chinese character reading of kindergarten children? *Infant and Child Development, 28*(3), e2122. https://doi.org/10.1002/icd.2122

Yin, B., & Rohsenow, J. S. (1994). *Modern Chinese characters*. Sinolingua.

Yu, B., Cao, H. (1992). An investigation of the effect of stroke number on the identification of Chinese characters and a discussion of the effect of stroke frequency. *Acta Psychologica Sinica, 24*(2), 120–126.

Zhang, Y., Zhang, L., Shu, H., Xi, J., Wu, H., Zhang, Y., & Li, P. (2012). Universality of categorical perception deficit in developmental dyslexia: an investigation of Mandarin Chinese tones. *Journal of Child Psychology and Psychiatry, 53*(8), 874–882. https://doi.org/10.1111/j.1469-7610.2012.02528.x

Zhang, Z., Yuan, Q., Liu, Z., Zhang, M., Wu, J., Lu, C., Ding, G., & Guo, T. (2021). The cortical organization of writing sequence: Evidence from observing Chinese characters in motion. *Brain Structure and Function, 226*(5), 1627–1639. https://doi.org/10.1007/s00429-021-02276-x

Zhou, X., & Marslen-Wilson, W. (1999). The nature of sublexical processing in reading chinese characters. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(4), 819–837.

Zhu, X. (1988). Analysis of cueing function of phonetic components in modern Chinese. In X. Yuan (Ed.), *Proceedings of the symposium on the Chinese language and characters* (pp. 85–99). Guang Ming Daily Press.

Ziegler, J. C., Perry, C., Ma-Wyatt, A., Ladner, D., & Schulte-Körne, G. (2003). Developmental dyslexia in different languages: Language-specific or universal? *Journal of Experimental Child Psychology, 86*(3), 169–193. https://doi.org/10.1016/S0022-0965(03)00139-5

Ziegler, J. C., Bertrand, D., Tóth, D., Csépe, V., Reis, A., Faísca, L., ... & Blomert, L. (2010). Orthographic depth and its impact on universal predictors of reading: A cross-language investigation. *Psychological science, 21*(4), 551–559.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.