Applied Psycholinguistics (2024), 1–22
doi:10.1017/S0142716424000249

CAMBRIDGE UNIVERSITY PRESS

ORIGINAL ARTICLE

A database of ambiguous Chinese characters with measures for meaning dominance and meaning balance

Huilin Chen, Xu Xu, Haiquan Li, Xinyue Yu, Ruting Pan and Zhaoyang Zhang

Shanghai Jiao Tong University, Shanghai, China
Corresponding author: Xu Xu; Email: xu2xu3@gmail.com

(Received 18 January 2024; revised 9 July 2024; accepted 26 July 2024)

Abstract
Chinese characters hold great potential to help inform and enrich psycholinguistic research on lexical ambiguity as a large portion of them are ambiguous in nature with meaning varying from context to context. This report presents a psycholinguistic database that contains over 2000 characters with normative measures for meaning dominance and meaning balance, that is, the relative frequency of each meaning associated with a target character and the degree of balance across the meanings of the character. The measurement process takes advantage of the fact that, in Chinese, generating words containing a target character is the most convenient way to specify and disambiguate character meanings. Character meanings stored in ordinary people's mental lexicon are identified based on the words, along with a small portion of meaning descriptions, listed by over 900 native speakers. The measures of meaning dominance and meaning balance for the characters are derived from computing the relative frequencies of the meanings. Potential research and practical applications of the database, as a valuable tool, to enhance our understanding of the acquisition, representation, and processing of ambiguous lexical items are discussed.

Keywords: Chinese character; lexical ambiguity; meaning balance; meaning dominance

A lexical form can carry multiple meanings, for example, *bank*. This phenomenon termed lexical ambiguity is prevalent across languages (e.g., Britton, 1978; Clemmons, 2008; Peng et al., 2003). In English, Britton (1978) estimated that some 44% of words in *the American Heritage Dictionary of the English Language* (1969) are ambiguous, while in Chinese, approximately 40% of characters included in *the Dictionary of Chinese Character Information* (1988) have two or more different meanings (Liu et al., 2007). Remarkably, though widespread, lexical ambiguity does not seem to hinder daily communications among native speakers, prompting decades of research into the processing of ambiguous words.

The multifaceted nature of lexical ambiguity

Initial studies on lexical ambiguity only differentiated between ambiguous and unambiguous words and reported a processing advantage for ambiguous words in lexical decision tasks, evidenced by faster response and higher accuracy (e.g., Azuma & van Orden, 1997; Borowsky & Masson, 1996; Hino & Lupker, 1996; Pexman & Lupker, 1999). As research advances, researchers have begun to notice the heterogeneity among ambiguous words, making a distinction between homonyms and polysemes based on the (un)relatedness of meanings associated with the same lexical form. For instance, Rodd et al. (2002) defined homonyms as ambiguous words with unrelated meanings (e.g., *bark* "the outer covering of a tree" or "the sound made by a dog") and polysemes as those with related senses (e.g., *twist* with the original sense of "the physical act of rotating" as in *twist the ankle* and the metaphorical sense of "a deliberate action to misrepresent or alter the facts" as in *twist the truth*). They consulted *the Wordsmyth-English Dictionary* (Parks et al., 1998) to distinguish meanings and senses, in which meanings were listed as distinct lexical entries and senses were grouped under a single entry. With the dictionary-based number of meanings (dNoM) and dictionary-based number of senses (dNoS) orthogonally manipulated, they found a facilitatory effect of multiple related senses but an inhibitory effect of multiple unrelated meanings, revealing the complexity of the lexical ambiguity phenomenon and its effect on word recognition. More recently, some researchers (Chen et al., 2024) have pointed out that even such a distinction can be an oversimplification of the phenomenon. Many of the homonyms among the stimuli of past research are in fact hybrids. For example, according to *Wordsmyth*, the most cited homonym in the literature, *bank*, has two unrelated meanings, each encompassing six related senses. Specifically, the first meaning entry comprises *a heap* (n.), *a slope* (n.), *the edge of a river* (n.), *to embank* (v.i.), *to pile* (v.i.), and *to form* (v.t.), while the second meaning entry comprises *a financial institution* (n.), *a gambling reserve* (n.), *a supply* (n.), *to engage with a bank* (v.i.), *to operate a bank* (v.i.), and *to deposit* (v.t.). Chen et al. subsequently present a database with two continuous measures, based on native speakers' subjective assessment, to quantify both the plurality of the meanings and the relatedness of the meanings carried by ambiguous lexical forms.

Another critical aspect of lexical ambiguity directly related to lexical processing is the dominance or relative frequencies of the meanings associated with an ambiguous lexical form, which is the focus of the current research. The dominance of meanings distinguishes between balanced and unbalanced ambiguous words, that is, ambiguous words with meanings of roughly equal levels of dominance (e.g., *pitcher* can mean both "thrower" and "vessel") and ambiguous words with one dominant meaning and one or more subordinate meanings (e.g., *port* with the dominant meaning "harbor" and the subordinate meaning "wine"). In a classic study, Duffy and colleagues (1988) monitored participants' eye movements as they read sentences composed of two reversible clauses, with one clause containing an ambiguous word and the other serving as the disambiguating context. When the contextual clause appeared after the clause containing the ambiguous word, balanced ambiguous words caused significantly longer gazing durations relative to unbalanced ambiguous words. In contrast, when the subordinate-biasing contextual

https://doi.org/10.1017/S0142716424000249 Published online by Cambridge University Press

clause appeared before the clause containing the ambiguous word, longer gazing durations were found for unbalanced ambiguous words, but not for balanced ones. Similarly, Tsang and Chen (2013a) compared the morphemic priming effect with balanced ambiguous characters (e.g., 乳 "suckling" or "milk") and unbalanced ones (e.g., 指 with the dominant meaning "finger" and the subordinate meaning "pointing") in the recognition of compound words. In the masked priming experiment, when the common character of the prime and target words took on distinct meanings, morphemic priming nonetheless occurred with unbalanced ambiguous characters, whereby the characters' subordinate meanings entailed by the prime words (e.g., 指 "pointing" as in the prime 指责 "censure") facilitated the activation of the characters' dominant meanings and thus the recognition of the target words (e.g., 指 "finger" as in 指纹 "fingerprint"), whereas morphemic priming was absent with balanced ambiguous characters. These findings suggest that ambiguous lexical forms with and without a dominant meaning undergo distinct processing. The assessment of relative meaning frequency and the identification of dominant versus subordinate meanings thus hold important implications for this line of research.

Unique features of Chinese characters

Chinese characters present a highly promising avenue for research on lexical ambiguity and its resolution. As mentioned, a significant proportion of Chinese characters have multiple meanings (e.g., Chen et al., 2024; Liu et al., 2007), which is in part attributable to the relatively simple Chinese phonetic system. With over 70% of disyllabic words (i.e., two-character words) in modern Chinese (McEnery & Xiao, 2004; Xu, 2017), characters often appear as word constituents, with their exact meanings determined by their accompanying characters. For instance, the character 花 signifies "flower" in 花草 "flowers and grass" and "to spend" in 花钱 "to spend money." Ambiguous Chinese characters thus make a convenient tool to explore the contextual effect on ambiguity resolution. For instance, using eye-tracking technology, Shen et al. (2018) investigated how the preview of the contextual characters influenced the reading of the entire compound words (e.g., 树林 "woods," 树胶 "resin," and 树敌 "to make enemies") containing unbalanced ambiguous characters (e.g., 树 with the dominant meaning "tree" and the subordinate meaning "to make"). In three conditions, the previewed contextual characters were chosen (1) to have an identical meaning of the compound with the ambiguous character taking on its dominant meaning (e.g., 林 "woods" previewed for 树林 "woods"), (2) to bias toward the dominant meaning of the ambiguous character (e.g., 胶 "gum" previewed for 树胶 "resin"), or (3) to bias toward the subordinate meaning of the ambiguous character (e.g., 敌 "enemy" previewed for 树敌 "to make enemies"). Prolonged gaze durations and go-past times on compounds were found in the subordinate condition compared to the dominant and the identical conditions, revealing the effects of the semantic composition of the constituent characters and meaning dominance of the ambiguous characters in the process of ambiguity resolution.

https://doi.org/10.1017/S0142716424000249 Published online by Cambridge University Press

In fact, there has been a long line of research taking advantage of character ambiguity to investigate how sub-lexical semantics influence the recognition of Chinese compound words (e.g. Huang & Lee, 2018; Shen et al., 2018; Tsang & Chen, 2013a; 2013b; Tsang & Chen, 2014; Zhao et al., 2017; Zhou et al., 1999). For instance, Zhou and colleagues (1999) manipulated stimulus onset asynchrony (SOA). They found facilitatory priming for word pairs with shared ambiguous characters conveying the same meaning, regardless of SOA durations (e.g., 华丽 "magnificent" – 华贵 "luxurious"). However, when the ambiguous character in the prime and target conveyed different meanings (e.g., 华侨 "overseas Chinese" – 华贵 "luxurious"), the priming effect was absent in the long SOA condition. Tsang and Chen (2013b) implemented an unmasked priming lexical decision task (Exp. 3) with unbalanced ambiguous characters (e.g., 月 with the dominant meaning "moon" and the subordinate meaning "month"). They found that the dominant primes (i.e., primes associated with the dominant meaning of ambiguous morphemes, such as 月蚀 "*moon* eclipse") facilitated the recognition of the dominant target (e.g., 月饼 "*moon* cake"), whereas subordinate primes (e.g., 月薪 "*monthly* pay") enhanced the processing of subordinate targets (e.g., 月历 "*monthly* calendar"). These findings indicate an early activation of character meanings in compound word processing and highlight the utility of ambiguous characters in elucidating the processing of compound words. Stimuli sampling of these studies mostly relied on small-scale normative ratings serving directly for the purposes of their respective studies, which may potentially cause duplication of effort and questionable comparability of research findings across studies. The relatively large-scale database provided by Chen et al. (2024) contains two ambiguity measures: the perceived number of meanings (pNoM) for more than 4,000 characters and the perceived relatedness of meanings (pRoM) for a subset of the characters. They found that both measures significantly predicted the efficiency of character processing, even after controlling for factors such as character frequency and age of acquisition. With a computational approach, Hsieh et al. (2024) reported a finding similar to the effect of pRoM, which they termed as the effect of semantic consistency of a character across all compounds containing the character.

Lastly, polyphones are relatively common in the Chinese language. Approximately 11% of commonly used ambiguous characters are estimated to be heterophonic homographs, meaning that they have different meanings corresponding to different pronunciations. (Sun, 1995). In some cases, both heterophonic ambiguity and homophonic ambiguity can be found within the same character. For example, the character 省 can mean "province" and "to save" with the same pronunciation/sheng3/, as well as "to reflect" with a different pronunciation/xing3/. These characters provide an opportunity to study the phonetic impact on word recognition (e.g., Tsang, 2021). For instance, Tsang (2021) conducted a masked priming lexical decision task with heterophonic characters in Cantonese Chinese. They observed the facilitative priming for prime-target pairs with shared ambiguous characters of the same pronunciation and meaning (e.g., 长短/coeng4dyun2/"length"-长远/coeng4yun5/"long-term"), but not for those with different pronunciations and meanings (e.g., 长官/zoeng2gun1/"senior official"-长远/coeng4yun5/"long-term"). Future research can make use of Chinese characters and further explore how phonetics may modulate the effects of meaning dominance revealed in earlier studies.

https://doi.org/10.1017/S0142716424000249 Published online by Cambridge University Press

Approaches to assess meaning dominance

Currently, there lacks an extensive database for the dominance of character meanings, despite its pivotal significance for research on the processing of characters and compound words. In contrast, meaning dominance norms for alphabetic languages are widely available, established over time, and collected through various methods (e.g., English: Armstrong et al., 2012; DeLong et al., 2023; Gawlick-Grendell & Woltz, 1994; Maciejewski & Klepousniotou, 2016; Twilley et al., 1994; see Gilbert & Rodd, 2022 for spoken words; Spanish: Armstrong et al., 2016; Hebrew: Peleg & Eviatar, 2008; etc.).

One of the most conventional approaches is via the task of free association, in which participants are asked to write down the first word associate or all word associates they can think of for a given ambiguous word (e.g., Brocher et al., 2018; Cramer, 1970; DeLong et al., 2023; Duffy et al., 1988; Gorfein et al., 1982; Nelson et al., 1980; Peleg & Eviatar, 2008; Twilley et al., 1994; Zempleni et al., 2007). The generated associates are then clustered to form meaning categories based on perceived semantic similarity among them or categorized according to dictionary definitions. For example, Twilley et al. (1994) asked independent raters to cluster first associates on the basis of at least one distinguishable semantic feature (Exp. 1), whereas Nelson et al. (1980) consulted *the American Heritage Dictionary of the American Language* (1973) to define meaning categories before asking four raters to classify word associates accordingly. Other approaches include classifying meaning definitions generated by participants (Eckstein et al., 2011; Gawlick-Grendell & Woltz, 1994; Gilhooly & Logie, 1980; Onifer & Swinney, 1981; Tsang & Chen, 2013a, 2013b; Warren et al., 1977), sentences generated by participants (Wollen et al., 1980), or sentence completions (Yates, 1978).

More recently, some researchers have proposed an alternative approach for measuring meaning dominance (e.g., Armstrong et al., 2012; Armstrong et al., 2016; Maciejewski & Klepousniotou, 2016). Armstrong and his colleagues (2012) designed a norming software, eDoM, which allowed researchers to present dictionary definitions of an ambiguous word on a screen. Participants were required to estimate the percentage of the time when the word was used to refer to a certain meaning. The average meaning percentage across participants stood for the relative meaning frequency. Using this software, Armstrong and his colleagues (2012) collected relative meaning frequencies for 544 English homonyms and demonstrated the predicative validity of their norms with analyses of lexical decision data from the English Lexicon Project (Balota et al., 2007) and data from their previous report (Armstrong & Plaut, 2011). Also utilizing the eDoM norming software, they gathered relative meaning frequencies for 578 homonyms in two Spanish dialects and suggested that as few as seven ratings were sufficient to establish a highly stable set of ratings (Armstrong et al., 2016). Maciejewski and Klepousniotou (2016) similarly collected relative meaning frequencies for 100 homonyms in British English.

The current study

The current study aims to construct a normative database that supplies rating-based measures of meaning dominance and meaning balance for 2,059 Chinese characters.

https://doi.org/10.1017/S0142716424000249 Published online by Cambridge University Press

As Chinese characters make an ideal candidate for research on lexical ambiguity, the availability of this database can be expected to facilitate not only future investigations into the effect and the resolution of lexical ambiguity on Chinese character and word processing but also cross-language comparisons of this universal phenomenon. Due to the following considerations, the current study opted for the task of asking participants to specify character meanings over estimating meaning percentages or generating free associates. First, while meaning percentage has been validated as a reliable and effective measure for meaning dominance, its collection process relies heavily on dictionaries, which can deviate from ordinary people's mental lexicon. To capture the mental representations of ambiguous characters, we recruited ordinary native speakers to list the meanings of each character presented to them based on their lexical knowledge. Second, compared with generating free associates, this meaning listing task largely precludes random and idiosyncratic responses that often appear in the free association task (Armstrong et al., 2012). Lastly, as indicated earlier, character ambiguity gets resolved in their word context. Therefore, asking participants to list words containing a target character is a convenient way to specify and disambiguate character meanings and to tap into ordinary people's understanding about the character.

Method

Participants

A total of 969 participants took part in the meaning listing task anonymously over the Internet. They received monetary compensation after completing the task. All participants were self-reported native speakers of Mandarin Chinese, who spent the first seven years in mainland China (e.g., Warriner et al., 2013; Xu et al., 2022). Responses from 14 participants were excluded from data analysis due to failing to follow the task instructions and providing irrelevant responses throughout the task. For instance, one participant responded with the word 名词 "noun" for all characters on the list. To retain as many valid responses as possible, we made no further participant-wise screening. The final data analysis included valid responses from 955 participants (mean age = 29.92±6.43). Figure 1 illustrates the demographic characteristics of these participants.

Character sample

A list of ambiguous characters was retrieved from a recently released database that reported the perceived number of meanings (pNoM) for over 4,000 common Chinese characters (Chen et al., 2024). The pNoM was collected by asking participants to indicate the number of meanings they could think of for each character on a 5-point numeric scale, with "0" denoting "no meaning" and "4" "more than three meanings." Of all 4,363 characters in the database, there were 2,710 characters with mean pNoM equal to or greater than 1. We then consulted *the Xinhua Dictionary* (2011) to double-check the semantic plurality of each character, which resulted in the removal of 675 characters with only one dictionary meaning. Further, we found that two characters, 丢 and 丟, were of different fonts of the same

https://doi.org/10.1017/S0142716424000249 Published online by Cambridge University Press

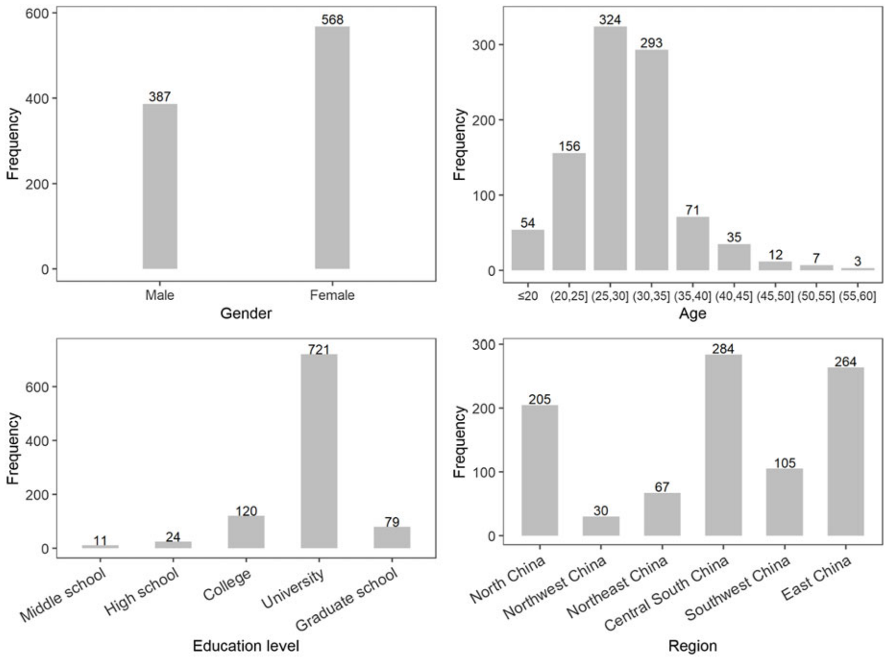

Figure 1. Gender, age, education level, and region distributions of the participants.

character and thus removed one of them from the sample. The remaining 2,034 characters, along with an additional 25 ambiguous characters from an ongoing project in our lab, constituted the final sample of 2,059 characters for this study. The characters were divided into 40 lists of approximately 50 characters, roughly matching on character frequency retrieved from SUBTLEX-CH (Cai & Brysbaert, 2010) and pNoM (Chen et al., 2024).

Instruction and procedure

Each participant was randomly assigned a list of approximately 50 characters. The instruction first stated that many Chinese characters possessed more than one meaning. Participants were asked to list all the meanings they could think of about each character. They were encouraged to respond, whenever possible, with words containing the target character to indicate and disambiguate the different meanings of the character. For the sake of succinctness, we termed these words containing the target character as word formations. When word formations were not readily available or apt to specify character meanings, they had the option to use phrasal or sentential descriptions. An example character 花 was provided with three words containing the character (花草, 花钱, and 眼花) to illustrate its different meanings ("flower," "to spend," and "blurry"). The instruction explicitly stated that the purpose of the study was to survey native speakers' understanding of character meanings, and the use of dictionaries was neither necessary nor appropriate. In case a character was unknown to the participants, they were instructed to respond with "N." In the end, they answered the demographic questions.

https://doi.org/10.1017/S0142716424000249 Published online by Cambridge University Press

Table 1. Coded responses, along with their raw frequencies, for the character 北

Word	Frequency	Tag	Word	Frequency	Tag
北方	10	M1	东北	1	M1
方向	3	M1	失败	1	M2
北面	3	M1	败北	1	M2
北边	2	M1	北京	8	P
南北	1	M1	北海	1	P
北门	1	M1	北大	1	P
方位方向	1	M1	北鼻	1	F
南的反义词	1	M1	贝壳	1	E

Note: M1: meaning 1-*north*; M2: meaning 2-*be defeated*; P: proper names-*Beijing*, *Beihai*, and *Beida*; F: phonologically based response-*baby*; E: error-*seashell*.

The responses, including word formations and brief meaning descriptions, listed for each character were first aggregated across participants for frequency tabulation. Table 1 presents an example character 北 with responses and their raw frequencies generated by 21 participants. Then, three undergraduate students and one graduate student, who were highly proficient native speakers of Mandarin Chinese, coded the responses to each character to categorize them into different meanings. The coding criteria were not given to the coders or extracted from the dictionary but established by the coders strictly according to their own lexical knowledge. Specifically, the three undergraduate students, who were blind to the ultimate objective of the study collectively coded the responses to an initial two lists of characters, based on which they discussed their coding criteria and finalized a coding scheme with four types of tags. First, the word formations and short meaning descriptions representing different meanings of a character were tagged with "M1," "M2," . . . , and "Mn." Second, the responses indicating that a character could serve as a proper name or part of a proper name were tagged with "P." Third, phonetically based responses were tagged with "F," which included (1) transliterated words, such as 巧克力 "chocolate," (2) Internet slang originated from intentional malapropism, such as 蓝瘦香菇 "blue thin mushroom" in place of 难受想哭 "feel sad and want to cry," and (3) onomatopoeia, such as 豁琅 (the sound of a teacup breaking). Lastly, obvious erroneous and irrelevant responses were tagged with "E," the various instances of which are outlined under *Data Screening* in the Results section. With the established coding scheme, the responses to each of the remaining 38 character lists received independent coding from two randomly paired coders. Coding discrepancies were resolved by a third coder.

Results

Data screening

Over 97% of characters received valid responses from 20 or more participants, while the rest of the characters received responses from no less than 17 participants. A total of 33,820 unique responses were collected, of which 1,273 responses (3.76%

of 33,820 unique responses) were tagged with "E" and removed from further analysis. These responses included (1) those that were obviously the word formations or the meaning interpretations of characters other than the target characters, such as characters highly orthographically similar to the target characters, for example, 夭 "young" mistaken for the character 天 "sky" and 曰 "speak" for the character 日 "sun"; (2) those that indicated only the parts of speech, for example, *a noun*, of the target characters, but failed to provide any specific meaning interpretations; and (3) those that were simply repetitions of the target characters, which were likely the results of failed attempts to find proper word formations. After data screening, within the responses to each character, those tagged with "P" (indicating the target character could appear as a proper name) and with "F" (indicating that the target character could serve as a phonetic symbol), respectively, were counted as two separate meaning categories in addition to the meaning categories tagged with "M1," "M2," ..., and "Mn."

To evaluate interrater reliability for meaning categorization, we calculated the percentage of consistent coding between the two coders for each of the 38 character lists, excluding the initial two-character lists coded collectively to establish the coding scheme. The percentage of coding consistency ranged from 75% to 94% across lists and from 79% to 91% across the six pairs of coders, with a mean of 84% ($SD = 5\%$).

Generated number of meanings (gNoM)

The total number of meaning categories associated with each character can be informative about perceived meaning dominance and dominance distribution of ambiguous characters. Intuitively, the greater the number of meaning categories identified for a character, the more likely these meanings are perceived to be relatively balanced in dominance. We therefore included in our database the total number of meanings generated (gNoM) by the participants. Figure 2 shows the gNoM distribution of 2059 characters, ranging from 1 to 8 (mean = 2.40; $SD = 1.13$). Characters possessing two or three distinguishable meanings constituted the majority of the characters, with two-meaning characters accounting for 37.35% and three-meaning characters 25.35% of the sample.

Similar to English dictionaries, Chinese dictionaries, such as *the Modern Chinese Dictionary* (2016), list unrelated *meaning*s as separate entries and group closely related *sense*s under the same entry. Past psycholinguistic research consequently tabulated both the number of meanings (dNoM) and the total number of senses (dNoS) in a dictionary compiled for a lexical form as indices for lexical ambiguity (e.g., Rodd et al., 2002). To compare gNoM with the numbers of dictionary meanings and senses, we counted the dNoM and dNoS, respectively, based on *the Modern Chinese Dictionary* (2016) for the characters in the current database. Correlation analysis and *t*-tests were conducted, with the measures (i.e., dNoM, dNoS, and gNoM) log-transformed to alleviate issues resulting from skewed distributions. The *t*-tests indicated that although it appeared that the meaning categories generated by ordinary native speakers were less specific or less comprehensive than the character senses listed in the dictionary ($t(2058) = -59.12$, $p < .001$), they seemed more fine-grained than the dictionary meaning

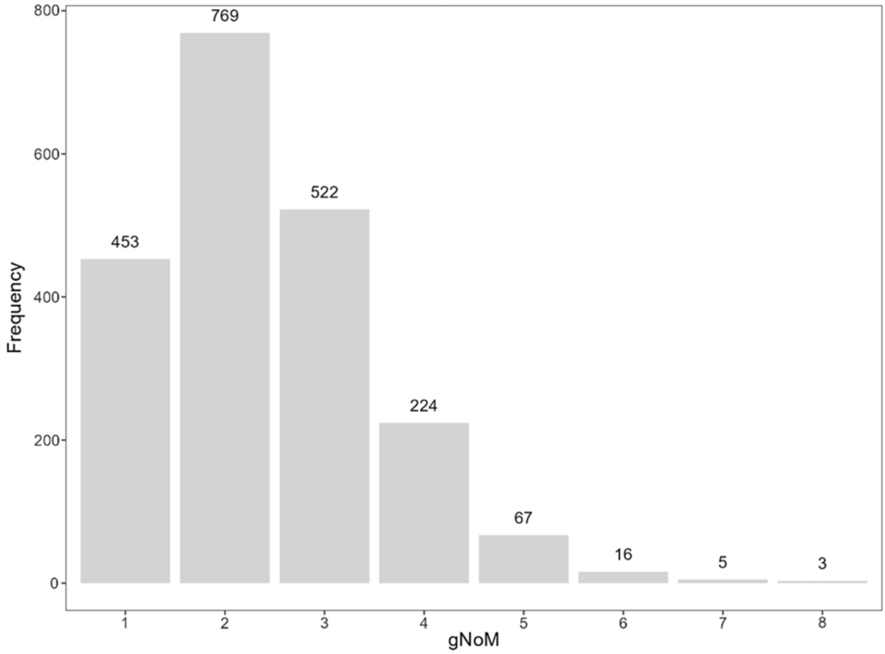

Figure 2. The distribution of generated number of meanings (gNoM).

entries associated with the characters ($t(2058) = 33.99$, $p < .001$). The correlation coefficients of gNoM with dNoS and dNoM were 0.456 and 0.326, respectively, demonstrating a closer alignment of gNoM with the number of dictionary senses relative to the number of dictionary meanings.

We also retrieved pNoM from the database of Chen et al. (2024) to compare with gNoM. The collection of pNoM and gNoM relied on the lexical knowledge of average native speakers but resorted to two different tasks, numeric rating versus meaning listing, respectively. In addition, the pNoM was computed by averaging the ratings across the participants and thus indicates, on average, how many meanings a participant could think of about a target character. As indicated earlier, for the purpose of assessing meaning dominance for a target character, gNoM was computed by aggregating all meanings across the participants. Therefore, we first computed on average how many meanings were listed by a participant for a target character and then compared the mean gNoM with pNoM. The rating-based pNoM is relatively more in line with the mean gNoM, $r = .325$, $p < .001$, than with gNoM, $r = .242$, $p < .001$. Further, the mean gNoM was significantly smaller than pNoM, $t(2058) = -26.474$, $p < .001$, reflecting the greater level of task difficulty of actually listing character meanings rather than simply counting the number of the meanings.

Assessment of meaning dominance and meaning balance

For each character, the relative frequency of a meaning was computed as the measure of the level of dominance of the meaning. Specifically, the number of

https://doi.org/10.1017/S0142716424000249 Published online by Cambridge University Press

Table 2. Measures of meaning balance (beta and D) and relative frequencies for the most and the second most dominant meanings (separate and combined) of the characters grouped by gNoM

gNoM	The most dominant meaning (*beta*)	The 2nd most dominant meaning	The most dominant and the 2nd most dominant meanings combined	*D*
1	1 (0)	0 (0)	1 (0)	1 (0)
2	.82 (.14)	.18 (.14)	1 (0)	.73 (.26)
3	.69 (.15)	.23 (.11)	.92 (.07)	.61 (.26)
4	.58 (.14)	.26 (.10)	.84 (.09)	.50 (.26)
5	.55 (.13)	.24 (.09)	.79 (.09)	.50 (.26)
6	.50 (.17)	.21 (.08)	.71 (.12)	.48 (.32)
7	.43 (.06)	.22 (.09)	.65 (.12)	.48 (.19)
8	.45 (.20)	.21 (.11)	.65 (.09)	.41 (.42)

Note: SDs are provided in parentheses.

responses under the same meaning category (i.e., P, F, M1, M2, …, or Mn) was divided by the total number of valid responses to the character. We ranked the meaning categories of each character based on their relative frequencies. Table 2 presents the means and standard deviations (*SD*s) of the relative frequencies associated with the most dominant meaning, the second most dominant meaning, and the first and second meanings combined for characters grouped by gNoM ranging from 1 to 8. It can be observed that as the gNoM increases, the relative frequencies of the most dominant and the second most dominant meanings decrease. However, regardless of the value of gNoM, the average combined relative frequency of the first and second most dominant meanings remained over 60%.

Given the prominence of the first and the second most dominant meanings of the characters, in the database, we provide both the measure *beta* and the measure *D* (Armstrong, et al., 2012) as the indices for meaning balance, that is, the degree of balanced versus unbalanced ambiguity of Chinese characters. The measure *beta* was simply the relative frequency of the most dominant meaning. That is, the greater the value of *beta* is, the more dominant the most prominent meaning is, and the more unbalanced the meanings of the ambiguous character are. In the extreme case where $beta = 1$, the character that carries only one meaning is considered to be unambiguous. Of the 2, 059 characters, their *beta* values ranged from 0.23 to 1, with 453 characters judged to be unambiguous (gNoM = 1, $beta = 1$).

The other measure *D* assesses meaning balance by quantifying the relative dominance of the first meaning over the second meaning, which was calculated with the formula of ([1st meaning frequency – 2nd meaning frequency]/1st meaning frequency) (Armstrong et al., 2012). The measure *D* took both of the most and the second most dominant meanings of a given character into consideration, thus providing a relatively more comprehensive estimate of meaning balance. The values of *D* for the 2,059 characters ranged from 0 to 1, which increased as the two most dominant meanings became more unbalanced. At one extreme, the first and the

https://doi.org/10.1017/S0142716424000249 Published online by Cambridge University Press

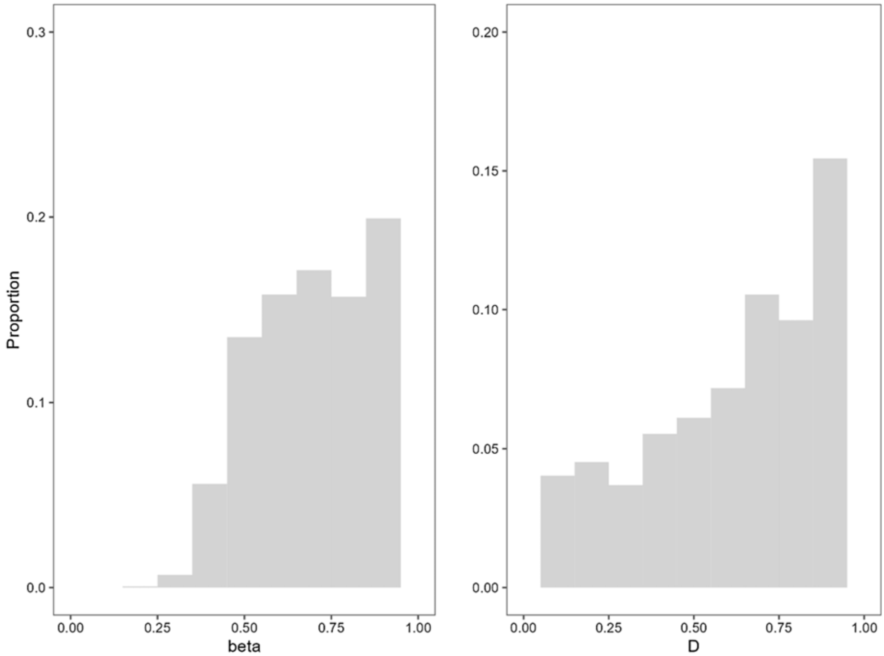

Figure 3. The distributions of beta and the D for characters with gNoM > 1.

second most dominant meanings of 21 characters were of the same relative frequency, and their *D* values equaled 0. At the other extreme, the 453 characters with only one meaning had a *D* value of 1. The correlation between *beta* and *D* was nevertheless very high ($r = .95$, $p < .001$), indicating similar efficacy of the two measures. Table 2 reports the means and *SD*s of these two measures for characters again grouped by gNoM.

Figure 3 shows the histograms of the *beta* and the *D* for 1,606 ambiguous characters with gNoM > 1, excluding the 453 characters with gNoM = 1. The distributions of both *beta* and *D* were moderately left-skewed, similar to the distribution for English ambiguous words (Armstrong et al., 2012). The left-skewed distributions suggested that the majority of ambiguous characters tended to be semantically unbalanced. Previous studies took different beta values as the cutoff point to classify balanced versus unbalanced ambiguous words (e.g., .75 in Armstrong et al., 2012 and Mirman et al., 2010; .65 in Armstrong & Plaut, 2011; .41–.59 for balanced ambiguous words and .69–.93 for unbalanced ones in Klepousniotou & Baum, 2007). Referring to these criteria, the current database has 58.81% (1,211 of 2,059) characters with *beta* values exceeding .75, 72.17% (1,486 of 2,059) characters exceeding .65, and 90.38% (1,861 of 2,059) characters exceeding .50. Even when excluding the 453 characters with gNoM = 1, we still observe that 47.2%, 64.32%, or 87.67% of the 1,606 characters should be considered unbalanced, with the cutoff beta value of.75, .65, or .50, respectively. That is, no matter what criteria were adopted to define unbalanced ambiguous characters, they tended to make up the majority of the sample. This finding emphasized the necessity to take relative meaning frequencies into consideration when constructing stimuli for experiments on ambiguous characters.

Relations of meaning balance measures with other variables

To inform future researchers, we also present here (Table 3) the correlations of the measures of meaning balance (i.e., *beta* and *D*) with other linguistic and psycholinguistic variables that are available and often included in research on Chinese characters, including the rating-based perceived number of meanings (pNoM; Chen et al., 2024) and perceived relatedness of meanings (pRoM; Chen et al., 2024), character frequency (CHARCount.log; SUBTLEX-CH by Cai & Brysbaert, 2010), age of acquisition (AoA; Cai et al., 2022), number of words formed by a character (NWF.log; MELD-SCH by Tsang et al., 2018), reaction time and error rate in character processing (zRT.Tsang and Error.log. Tsang; MELD-SCH by Tsang et al., 2018; zRT.Sze and Acc. Sze; CLP by Sze et al., 2014[1]). Behavioral data from two databases with different character coverages, that is, MELD-SCH (Tsang et al., 2018) and CLP (Sze et al., 2014), were incorporated to comprehensively evaluate the impact of meaning dominance on character processing. MELD-SCH represents the latest database involving simplified characters, while CLP shares more characters with the current database.

Both *beta* and *D* showed relatively strong positive correlations with the measure for relatedness of a character's meanings (pRoM). Greater pRoM was associated with lower meaning balance (i.e., greater dominance of the most prominent meaning). Conspicuously, highly related senses of a character tend to be perceived as the same meaning by many native speakers and listed under the same meaning entry in the dictionary by lexicographers. This broad meaning of a character that can be instantiated and further specified in a diverse range of lexical contexts therefore appears more likely to be perceived as dominant. The associations of both *beta* and *D* with the other variables appeared trivial, with magnitudes of the correlations consistently much lower or nonsignificant, including the variables representing character processing efficiency, that is, zRT, error rate, and accuracy rate. We discuss these findings in the next section.

Discussion

The current study assessed meaning dominance and meaning balance for 2,059 ambiguous Chinese characters, which are vital to research on the representation and processing of lexical ambiguity. We provide in the database two measures, *beta* and *D*, for meaning balance of an ambiguous character, along with the relative frequency (i.e., dominance) of each meaning of the character and gNoM (i.e., the number of meanings generated for the character by native speakers). Similar to reports on other languages (Armstrong et al., 2012; 2016), the distributions of *beta* and *D* in the current study indicate that the majority of ambiguous characters are in fact unbalanced with one or two meanings dominating the other(s). Given that balanced ambiguous lexical forms are of particular interest to ambiguity research, these findings highlight the need to consult balance measures for the purpose of stimuli construction when investigating the processing and resolution of lexical ambiguity.

Table 3. Correlations between beta, D, and other variables

	1	2	3	4	5	6	7	8	9	10
1. beta										
2. D	.946** (2059)									
3. pNoM	−.205** (2059)	−.162** (2059)								
4. pRoM	.594** (948)	.505** (948)	−.233** (1052)							
5. CHAR Count.log	−.190** (2058)	−.133** (2058)	.622** (4274)	−.163** (1052)						
6. AoA	.119** (1970)	.079** (1970)	−.567** (3715)	.066* (1041)	−.680** (3667)					
7. NWF.log	−.175** (416)	−.114* (416)	.689** (870)	−.021 (207)	.853** (870)	−.640** (692)				
8. zRT.Tsang	.044 (416)	.014 (416)	−.662** (870)	−.223** (207)	−.748** (870)	.620** (692)	−.742** (870)			
9. Error.log. Tsang	.057 (416)	.021 (416)	−.636** (870)	.008 (207)	−.676** (870)	.527** (692)	−.677** (870)	.784** (870)		
10. zRT.Sze	.118** (1450)	.084** (1450)	−.516** (2483)	.140** (786)	−.678** (2470)	.572** (2288)	−.623** (381)	.714** (381)	.541** (381)	
11. Acc. Sze	−.048 (1450)	−.044 (1450)	.353** (2483)	.005 (786)	.446** (2470)	−.338** (2288)	.338** (381	−.524** (381)	−.403** (381)	−.699** (2483)

Note: * $p < .005$, ** $p < .001$. Ns are presented in parentheses.

https://doi.org/10.1017/S0142716424000249 Published online by Cambridge University Press

Analyses of the reported measures revealed that *beta* and *D* appeared to be equally effective in assessing meaning balance for the characters included in this database. In addition, they both showed positive correlations with the degree to which the multiple meanings of a character were perceived to be closely related to one another (pRoM). Further, *beta* and *D* appeared to exert subtle effects on the processing of Chinese characters.

Measures in the current database

The database reported in the current study includes a few indices through which researchers may gain insight into the meaning representations of Chinese ambiguous characters. First, gNoM represents the total number of meanings generated by ordinary native speakers. They listed word formations containing a target character or brief descriptions to represent the meaning(s) of the target character. We included in the database only ambiguous characters (both dNoM and pNoM > 1) but found a total of 453 characters with gNoM $= 1$. The value of dNoM represents the total number of dictionary meanings, and the value of pNoM represents the average number of meanings rated by a group of native speakers on a 5-point numeric scale. The task to collect gNoM however required a group of native speakers to provide written responses about the character meanings, which was conspicuously more challenging. The value of gNoM $= 1$ therefore likely reflects the fact that these 453 characters each have an extremely dominant meaning, rather than only one meaning, and the subordinate meaning(s) may be too difficult to verbalize or too trivial to substantiate. The distribution of gNoM indicates that almost two-thirds of the characters were associated with two or three meanings. About 15% were found to have more than three meanings.

As presented in the Results section, we compared the values of gNoM with both dNoM (the number of broad meaning entries for a character in a dictionary) and dNoS (the total number of specific senses listed under all meaning entries for a character in a dictionary). Its stronger association with dNoS indicates that the total number of meanings "complied" by ordinary native speakers, that is, the gNoM, by and large, reflects the amount of specific dictionary senses compiled by professional lexicographers. Furthermore, *t*-tests indicate that although understandably meanings generated by ordinary people were not as specific or comprehensive as the senses listed in a dictionary, they were more fine-grained than the broad dictionary meaning entries of the characters.

It is worth noting that although gNoM and pNoM are both based on the lexical knowledge of native speakers, they exhibit only a moderate correlation. As suggested earlier, this should be attributed to the fact that the meaning rating task to collect pNoM and the meaning listing task to collect gNoM tapped into the two different aspects of one's language capacity, that is, comprehension versus production. However, both pNoM and gNoM demonstrated greater alignment with dictionary senses (dNoS) relative to dictionary meanings (dNoM). In the current study, we reported a stronger correlation of gNoM with dNoS ($r = .456$) than with dNoM ($r = .326$), while Chen et al. (2024) reported a stronger correlation of pNoM with dNoS ($r = .613$) than with dNoM ($r = .225$). These findings consistently indicate that, relatively speaking, the representation of Chinese character semantics in an individual's mental lexicon mirrors the fine-grained character senses, rather than coarsely classified character meanings, in the dictionary that documents the collective knowledge of people in the same language community.

The primary contribution of the database is the provision of two measures for meaning balance, that is, *beta* and *D*, which are included along with the relative frequencies associated with a character's meanings to compute these measures. To our knowledge, this is the first meaning dominance and meaning balance database for a relatively large set of ambiguous Chinese characters, which holds significant implications for future research on the semantic representation and processing of Chinese characters. Regardless of whether the second most dominant meaning of a character was taken into account, the two measures of meaning balance, *beta* and *D*, demonstrated a strong correlation ($r = .95$) and showed a similar left-skewed distribution. That is, unbalanced ambiguous characters substantially outnumber balanced ambiguous characters, similar to the findings in alphabetic languages, such as English (e.g., Armstrong et al., 2012) and Spanish (e.g., Armstrong et al., 2016). Thus, a random sampling of stimuli from any of these databases is likely to result in a sample containing more unbalanced items than balanced ones, indicating the necessity to constrain or stratify, depending on the purpose of the study, the sampling process with measures for meaning balance.

In addition, Table 3 shows that both *beta* and *D* exhibit significant positive correlations with the relatedness of character meanings (*beta* and pRoM: $r = .594$; *D* and pRoM: $r = .505$). That is, the more unbalanced an ambiguous character is, the more likely its meanings are perceived to be related to one another. This certainly does not mean that the dominant meaning and the subordinate meaning of an unbalanced homonymous character (e.g., 树 with the dominant meaning *tree* and the subordinate meaning *to make*) are more likely to be perceived as related to each other. More plausibly, the positive correlations indicate that some of the multiple senses of a character, if highly related (i.e., of high degrees of semantic overlap), tend to be perceived as one dominant meaning, and thus the character is more likely to be judged as an unbalanced ambiguous lexical form. The recent computational study by Hsieh et al. (2024) appears to support this argument. They quantified the semantic consistency of a character by averaging cosine distances between the character and all words that contain the character. That is, instead of categorizing and assessing broad character meanings, their approach directly measured word formations carrying the subtle differences of character senses. They reported a database including 5,188 traditional Chinese characters. Despite the script difference, we found that their position-general, type-based metric of semantic consistency showed a positive correlation with pRoM ($r = .347$, $p < .001$, $N = 951$) and further positive correlations with the dominance of character meanings (*beta*: $r = .284$, $p < .001$, $N = 1868$; *D*: $r = .230$, $p < .001$, $N = 1868$). Taken together, these findings indicate that perceived dominance of a character meaning is to some degree related to how frequently the meaning is instantiated in diverse lexical contexts.

The impacts of meaning dominance on character processing

Given the relatively strong positive correlation of *beta* and *D* with the relatedness of meanings (i.e., pRoM), we suspect that the effect of meaning relatedness on the

https://doi.org/10.1017/S0142716424000249 Published online by Cambridge University Press

processing of ambiguous words, as reported in previous studies (e.g., Armstrong & Plaut, 2008; 2011; Beretta et al., 2005; Rodd et al., 2002; Tamminen et al., 2006), may be partially attributable to the influence of meaning dominance. As is reviewed, Rodd et al. (2002) manipulated dNoM and dNoS in a 2 × 2 factorial design to investigate the interaction between the number of unrelated meanings (many vs. few) and the number of related senses (many vs. few). They reported a sense advantage but a meaning disadvantage in a lexical decision task. We randomly sampled 40 characters from the current database, 10 from each of the four categories by crossing few versus many unrelated meanings and few versus many related senses according to the *Modern Chinese Dictionary* (2016), with matched character frequency (Cai & Brysbaert, 2010). Across conditions, characters with few versus many unrelated meanings did not demonstrate significant differences in *beta* ($F(1,38) = .887, p = .352$) or *D* ($F(1,38) = .151, p = .700$). In contrast, characters categorized by few versus many related senses exhibited noteworthy differences in both *beta* ($F(1,38) = 4.889$, $p = .033$) and *D* ($F(1,38) = 3.283$, $p = .078$). This implies that the observed processing advantage of related senses might not have stemmed solely from the semantic overlap between senses and could have also been explained by the dominance of a focal sense among the related senses. Alternatively, the perceived relative dominance of a focal sense might be the result of semantic overlap among these related senses due to cumulative simultaneous activations of these senses. Specifically, as indicated by the above analysis on the computational metric of semantic consistency (Hsieh et al., 2024), a character's semantic consistency across its word formations is positively related to how dominant the character's most salient meaning is as perceived by native speakers. The magnitude of this correlation is moderate, which however might be the result of script (traditional vs. simplified) and measurement (computational vs. behavioral) differences between Hsieh et al. and the current study.

Both *beta* and *D* were found to be significantly correlated with reaction time (zRT) in a lexical decision task retrieved from the CLP (Sze et al., 2014). However, these correlation coefficients were modest. When employing the forward method to construct a regression model including *beta* along with other variables previously demonstrated to impact Chinese character processing (i.e., CharCount.log, AoA, pNoM, pRoM) to predict zRT, *beta* did not emerge as a significant predictor. The subtle impact of meaning dominance on character processing appears to have diminished with reduced sample size and decreased character frequency, as evidenced by the nonsignificant correlations of both *beta* and *D* with zRT and error rate from the MELD-SCH database (Tsang et al., 2018). Tsang and colleagues (2018) deliberately sampled low-frequency characters that were not included in the CLP (Sze et al., 2014) when constructing the MELD-SCH database. These findings suggest that the dominance effect on character processing may be specific to certain cases with a conglomerate of variables that come into play, including character frequency, relatedness of meanings, and possibly many others (e.g., Chen et al., 2024; Sze et al., 2014; Tsang et al., 2018). Experimental studies relying on factorial designs with powerful manipulation may be able to more effectively detect this effect on the online processing of characters. The findings of the current study however reveal meaning dominance as an important dimension of character representation in native speakers' mental lexicon.

Potential uses of the database

As indicated above, the construction of experimental stimuli can be better guided and tailored for the specific goals of research studies. First, the measures of *beta* and *D* can be used to distinguish between balanced and unbalanced ambiguous characters. Whereas previous studies suggest that balanced and unbalanced ambiguous words elicit distinct response patterns (e.g., Duffy et al., 1988; Armstrong et al., 2012), the correlation analysis in the current study indicates a subtle impact of meaning dominance on the processing of characters. With the help of more precise sampling, future research can explore the boundary conditions of the ambiguity effects on character processing. Second, studies focusing on ambiguity resolution (e.g., Atchley & Kwasny, 2003; Bitan et al., 2017; Bilenko et al., 2009; Klepousniotou et al., 2012) often require identification of the dominant versus subordinate meanings of an ambiguous word and quantification of their difference in relative dominance. The current database provides this necessary information for simplified Chinese characters. Specifically, relative frequency is provided, as the index for relative dominance, for each meaning associated with the target character. Third, as instructed, the participants provided word formations of target characters to specify and disambiguate character meanings. These word formations along with their respective frequencies indicate the typical lexical contexts in which the target characters are likely to appear. Studies on ambiguity resolution with Chinese characters often adopt lexical contexts to specify a certain meaning of an ambiguous character (e.g., Tsang & Chen, 2013a; 2013b; Wu et al., 2017; Wu et al., 2020), and the list of word formations included in this database is therefore a convenient tool for this line of research.

In addition to the ambiguity research, this database of Chinese characters can also be a useful tool to study other semantic dimensions of Chinese characters and words. For instance, Peng et al. (2024) collected sentiment annotations for 3,827 simplified Chinese characters. Assuming that their participants would provide valence ratings based on the most dominant meaning of an ambiguous character, we obtained character valence ratings from the database of Peng et al. (2024) and retrieved valence ratings (Xu et al., 2022) for the most dominant word formations in the current database. A strong correlation ($r = 0.82$, $p < .001$, N = 1301) was indeed found between character valance and the valence of its most frequently listed word formation. Such finding also attests to the efficacy of using word formations to study semantic representations of Chinese characters.

Finally, the current database also holds pedagogical implications. The acquisition of ambiguous vocabulary poses a challenge for native speakers as well as learners of Chinese as a foreign language, given the complex mapping between form and meaning. In the teaching of ambiguous Chinese characters, instructors can embed them within multisyllabic words guided by relative meaning dominance, thereby enhancing instructional efficacy. The words elicited by characters and relative frequencies of specific meanings in the current database are highly applicable in this context.

Methodological considerations and limitations

The participants in the current study were asked to use word formations, as much as possible, to denote character meanings. This methodological decision was

https://doi.org/10.1017/S0142716424000249 Published online by Cambridge University Press

primarily attributable to the fact that modern Chinese is bimorphemic in nature with over 70% of words being two-character words (McEnery & Xiao, 2004). Although many characters can stand alone and be considered single-character words, they frequently appear as constituents of two- or multicharacter words which themselves are therefore perfect tools to annotate and disambiguate the meanings of their constituent characters. Similar research on alphabetic languages (e.g., Twilley et al., 1994) asked participants to generate free associates of the target words, which were then categorized based on dictionary definitions of the target words. This approach can be effective for homonyms with clearly different and unrelated meanings but challenging for polysemes with related or less well-delineated meanings, a drawback long recognized by past researchers (e.g., Armstrong et al., 2012). Categorizing word formations (and brief meaning descriptions) appeared to be a much more straightforward task, the highly proficient coders in the current study thus demonstrated high levels of coding consistency.

However, same as any judgment tasks based on coders' knowledge and experiences, meaning categorization is inevitably subjective. Although the four coders came from different regions of the country, they shared similar demographic characteristics: young adults studying in a large university for academic degrees. Further, through a collectively determined coding scheme, they were able to reach high levels of agreement. Therefore, their categorizations of the word formations and meaning descriptions might be specific to this demographic group. Thus, it might be meaningful for future research to explore and compare lexical knowledge and experiences of other subpopulations, for example, people who did not receive formal secondary education and perhaps more importantly clinical populations, against the norms provided in this database. Similarly, the study was conducted online to collect word formations and meaning descriptions. On the one hand, it had the potential to recruit participants from a wide range of geographical regions. On the other hand, like most online studies, college students and young adults constituted the dominant majority of the participants. Such a sample afforded the current database a wealth of popular expressions shared among young, educated adults and made it potentially more suitable for linguistic and psycholinguistic research on this population.

Replication package. The reported database, including study materials (i.e., target characters), data (i.e., coded responses), and computed measures, as well as analysis codes (i.e., SPSS syntax for inferential analyses), is available at https://osf.io/yr6qz/.

Acknowledgments. The authors wish to thank all participants for their valuable contributions to this study. We would also like to thank three anonymous reviewers for their detailed comments and helpful suggestions on the choice of terminology, analysis of the data, and discussion of the results, which helped to improve this manuscript.

Competing interests. The authors declare none.

Funding. This study was supported by Shanghai Jiao Tong University (grant no. WF220414005).

Note

1 We thank an anonymous reviewer for their suggestion to incorporate the CLP in the analysis.

https://doi.org/10.1017/S0142716424000249 Published online by Cambridge University Press

References

Armstrong, B. C., & Plaut, D. C. (2008). Settling dynamics in distributed networks explain task differences in semantic ambiguity effects: Computational and behavioral evidence. *Proceedings of the Annual Meeting of the Cognitive Science Society*, (January), 273–278.

Armstrong, B. C., & Plaut, D. C. (2011). Inducing homonymy effects via stimulus quality and (not) nonword difficulty: Implications for models of semantic ambiguity and word recognition. In L. Carlson, C. Hölscher, & T. F. Shipley (Eds.), *Expanding the space of cognitive science: Proceedings of the 33rd Annual Meeting of the Cognitive Science Society* (pp. 2223–2228). Cognitive Science Society.

Armstrong, B. C., Tokowicz, N., & Plaut, D. C. (2012). eDom: norming software and relative meaning frequencies for 544 English homonyms. *Behavior Research Methods*, **44**(4), 1015–1027.

Armstrong, B. C., Zugarramurdi, C., Cabana, Á., Valle Lisboa, J., & Plaut, D. C. (2016). Relative meaning frequencies for 578 homonyms in two Spanish dialects: a cross-linguistic extension of the English eDom norms. *Behavior Research Methods*, **48**(3), 950–962.

Atchley, R. A., & Kwasny, K. M. (2003). Using event-related potentials to examine hemispheric differences in semantic processing. *Brain and Cognition*, **53**(2), 133–138.

Azuma, T., & Van Orden, G. C. (1997). Why SAFE is better than FAST: the relatedness of a word's meanings affects lexical decision times. *Journal of Memory and Language*, **36**(4), 484–504.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., & Treiman, R. (2007). The English lexicon project. *Behavior Research Methods*, **39**, 445–459.

Beretta, A., Fiorentino, R., & Poeppel, D. (2005). The effects of homonymy and polysemy on lexical access: an MEG study. *Cognitive Brain Research*, **24**(1), 57–65.

Bilenko, N., Grindrod, C., Myers, E., & Blumstein, S. (2009). Neural correlates of semantic competition during processing of ambiguous words. *Journal of Cognitive Neuroscience*, **21**(5), 960–975.

Bitan, T., Kaftory, A., Meiri-Leib, A., Eviatar, Z., & Peleg, O. (2017). Phonological ambiguity modulates resolution of semantic ambiguity during reading: an fMRI study of Hebrew. *Neuropsychology*, **31**(7), 759–777.

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.

Britton, B. K. (1978). Lexical ambiguity of words used in English text. *Behavior Research Methods and Instrumentation*, **10**, 1–7.

Brocher, A., Koenig, J. P., Mauner, G., & Foraker, S. (2018). About sharing and commitment: the retrieval of biased and balanced irregular polysemes. *Language, Cognition and Neuroscience*, **33**(4), 443–466.

Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PLOS One*, **5**(6), e10729.

Cai, Z. G., Huang, S., Xu, Z., & Zhao, N. (2022). Objective ages of acquisition for 3300+ simplified Chinese characters. *Behavior Research Methods*, **54**, 311–323.

Chen, H., Xu, X., & Wang, T. (2024). Assessing lexical ambiguity of simplified Chinese characters: plurality and relatedness of character meanings. *Quarterly Journal of Experimental Psychology*, **77**(4), 677–693.

Clemmons, K. (2008). The Problem of Polysemy in the First Thousand Words of the General Service List: A Corpus Study of Secondary Chemistry Texts. Dissertations & Theses Gradworks.

Cramer, P. (1970). A study of homographs. In L. Postman & G. Keppel (Eds.), *Norms of word association* (pp. 361–370). Academic Press.

DeLong, K. A., Trott, S., & Kutas, M. (2023). Offline dominance and zeugmatic similarity normings of variably ambiguous words assessed against a neural language model (BERT). *Behavior Research Methods*, **55**, 1537–1557.

Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language*, **27**(4), 429–446.

Eckstein, D., Kubat, M., & Perrig, W. J. (2011). Visible homonyms are ambiguous, subliminal homonyms are not: a close look at priming. *Consciousness and Cognition*, **20**(4), 1327–1343.

Gawlick-Grendell, L. A., & Woltz, D. J. (1994). Meaning dominance norms for 120 homographs. *Behavior Research Methods, Instruments, & Computers*, **26**(1), 5–25.

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.

Gilbert, R. A., & Rodd, J. M. (2022). Dominance Norms and Data for Spoken Ambiguous Words in British English. *Journal of Cognition*, **5**(1), 1–14.

Gorfein, D. S., Viviani, J. M., & Leddo, J. (1982). Norms as a tool for the study of homography. *Memory & Cognition*, **10**(5), 503–509.

Hino, Y., & Lupker, S. J. (1996). Effects of polysemy in lexical decision and naming: an alternative to lexical access accounts. *Journal of Experimental Psychology: Human Perception and Performance*, **22**(6), 1331–1356.

Hsieh, C.-Y., Marelli, M., & Rastle, K. (2024). Beyond quantity of experience: exploring the role of semantic consistency in Chinese character knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **50**(5), 819–832.

Huang, H. W., & Lee, C. Y. (2018). Number of meanings and number of senses: an ERP study of sublexical ambiguities in reading Chinese disyllabic compounds. *Frontiers in Psychology*, **9**, 324–332.

Institute of Linguistics, Chinese Academy of Social Sciences. (2011). *Xīnhuá Dictionary* (11th edition). The Commercial Press.

Institute of Linguistics, Chinese Academy of Social Sciences. (2016). *Xiàndài Hànyǔ Cídiǎn* (7th edition) [*Modern Chinese Dictionary*]. The Commercial Press.

Klepousniotou, E., & Baum, S. R. (2007). Disambiguating the ambiguity advantage effect in word recognition: an advantage for polysemous but not homonymous words. *Journal of Neurolinguistics*, **20**(1), 1–24.

Klepousniotou, E., Pike, G. B., Steinhauer, K., & Gracco, V. (2012). Not all ambiguous words are created equal: an EEG investigation of homonymy and polysemy. *Brain and Language*, **123**(1), 11–21.

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

Maciejewski, G., & Klepousniotou, E. (2016). Relative Meaning Frequencies for 100 Homonyms: British eDom Norms. *Journal of Open Psychology Data*, **4**, 6–10.

McEnery A., & Xiao Z. (2004). The Lancaster Corpus of Mandarin Chinese: A Corpus for Monolingual and Contrastive Language Study. *Proceedings of the 4th International Conference on Language Resources and Evaluation, LREC* 2004, 1175–1178.

Mirman, D., Strauss, T., Dixon, J., & Magnuson, J. (2010). Effect of representational distance between meanings on recognition of ambiguous spoken words. *Cognitive Science*, **34**(1), 161–173.

Morris, W. (1969). *The American Heritage Dictionary of the English Language* (1st edition). Houghton Mifflin.

Morris, W. (1973). *The American Heritage Dictionary of the English Language* (2nd edition). Houghton Mifflin.

Nelson, D. L., McEvoy, C. L., Walling, J. R., & Wheeler, J. W. (1980). The University of South Florida homograph norms. *Behavior Research Methods & Instrumentation*, **12**(1), 16–37.

Onifer, W., & Swinney, D. A. (1981). Accessing lexical ambiguities during sentence comprehension: effects of frequency of meaning and contextual bias. *Memory & Cognition*, **9**(3), 225–236.

Parks, R., Ray, J., & Bland, S. (1998). *Wordsmyth English dictionary—Thesaurus* [ONLINE], Available: http://www.wordsmyth.net/, University of Chicago.

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

Peng, C., Xu, X., & Bao, Z. (2024). Sentiment annotations for 3827 simplified Chinese characters. *Behavior Research Methods*. **56**(2), 651–666.

Peleg, O., & Eviatar, Z. (2008). Hemispheric sensitivities to lexical and contextual information: evidence from lexical ambiguity resolution. *Brain and Language*, **105**(2), 71–82.

Pexman, P. M., & Lupker, S. J. (1999). Ambiguity and visual word recognition: can feedback explain both homophone and polysemy effects? *Canadian Journal of Experimental Psychology*, **53**, 323–334.

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.

Science Publishers (1988). *Dictionary of Chinese Character Information*. Science Publishers.

Shen, W., Li, X., & Pollatsek, A. (2018). The processing of Chinese compound words with ambiguous morphemes in sentence context. *Quarterly Journal of Experimental Psychology*, **71**, 131–139.

Sun, Y. S. (1995). Xiǎoyì xiàndài hànyǔ chángyòngzì zhōng de duōyīnduōyìzì [Brief discussion on polyphonic and polysemous characters among common characters in modern Chinese]. *Jíníng Shīfàn Dàxué Xuébào [Journal of Jining Normal University].* **2**, 53–55.

Sze, W. P., Liow, S. J. R., & 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.

Tamminen, J., Cleland, A. A., Quinlan, P. T., & Gaskell, M. G. (2006). Processing semantic ambiguity: Different loci for meanings and senses. *Proceedings of the 28th Annual Conference of the Cognitive Science Society*, (January), 2222–2227.

Tsang, Y. K., & Chen, H. C. (2013a). Morpho-semantic processing in word recognition: evidence from balanced and biased ambiguous morphemes. *Journal of Experimental Psychology: Learning Memory and Cognition*, **39**(6), 1990–2001.

Tsang, Y. K., & Chen, H. C. (2013b). Early morphological processing is sensitive to morphemic meanings: evidence from processing ambiguous morphemes. *Journal of Memory and Language*, **68**(3), 223–239.

Tsang, Y. K., & Chen, H. C. (2014). Activation of morphemic meanings in processing opaque words. *Psychonomic Bulletin and Review*, **21**(5), 1281–1286.

Tsang, Y., Huang, J., Lui, M., Xue, M., Chan, Y. F., Wang, S., & Chen, H. (2018). MELD-SCH: a megastudy of lexical decision in simplified Chinese. *Behavioral Research Method*, 1763–1777.

Tsang, Y. K. (2021). Morphophonological activation in Chinese word recognition: evidence from heteronymic characters. *The Mental Lexicon*, **16**(2–3), 240–270.

Twilley, L. C., Dixon, P., Taylor, D., & Clark, K. (1994). University of Alberta norms of relative meaning frequency for 566 homographs. *Memory & Cognition*, **22**(1), 111–126.

Warren, R. E., Bresnick, J. H., & Green, J. P. (1977). Definitional dominance distributions for 20 English homographs. *Bulletin of the Psychonomic Society*, **10**(3), 229–231.

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, **45**(4), 1191–1207.

Wollen, K. A., Cox, S. D., Coahran, M. M., Shea, D. S., & Kirby, R. F. (1980). Frequency of occurrence and concreteness ratings of homograph meanings. *Behavior Research Methods & Instrumentation*, **12**(1), 8–15.

Wu Y., Tsang Y.-K., Wong A.-K., Chen H.-C. (2017). The processing of homographic morphemes in Chinese: an ERP study. *Neuroscience*, **32**(1),102–116.

Wu, Y., Duan, R., Zhao, S., & Tsang, Y. K. (2020). Processing ambiguous morphemes in Chinese compound word recognition: behavioral and ERP evidence. *Neuroscience*, **446**, 249–260.

Xu J. (2017). Text of Recent Chinese (ToRCH) 2014. Retrieved November 23, 2019, from http://corpus.bfsu.edu.cn/CORPORA.htm.

Xu, X., Li, J., & Chen, H. L. (2022). Valence and arousal ratings for 11,310 simplified Chinese words. *Behavior Research Methods*, **54**, 26–41.

Yates, J. (1978). Priming dominant and unusual senses of ambiguous words. *Memory & Cognition*, **6**(6), 636–643.

Zempleni, M. Z., Renken, R., Hoeks, J. C. J., Hoogduin, J. M., & Stowe, L. A. (2007). Semantic ambiguity processing in sentence context: evidence from event-related fMRI. *NeuroImage*, **34**(3), 1270–1279.

Zhao, S., Wu, Y., Li, T., & Guo, Q. (2017). Morpho-semantic processing in Chinese word recognition: an ERP study. *Acta Psychologica Sinica*, **49**(3), 296–306.

Zhou, X., Marslen-Wilson, W., Taft, M., & Shu, H. (1999). Morphology, orthography, and phonology reading Chinese compound words. *Language and Cognitive Processes*, **14**, 525–565.

Cite this article: Chen, H., Xu, X., Li, H., Yu, X., Pan, R., & Zhang, Z. (2024). A database of ambiguous Chinese characters with measures for meaning dominance and meaning balance. *Applied Psycholinguistics*. https://doi.org/10.1017/S0142716424000249