Behavior Research Methods (2021) 53:1697–1709
https://doi.org/10.3758/s13428-020-01525-x

A Chinese Conceptual Semantic Feature Dataset (CCFD)

Yaling Deng[1,2] · Ye Wang[1,2] · Chenyang Qiu[2] · Zhenchao Hu[3] · Wenyang Sun[4] · Yanzhu Gong[2] · Xue Zhao[5] · Wei He[5] · Lihong Cao[1,2,6]

Accepted: 12 December 2020 / Published online: 2 February 2021

Abstract

Memory and language are important high-level cognitive functions of humans, and the study of conceptual representation of the human brain is a key approach to reveal the principles of cognition. However, this research is often constrained by the availability of stimulus materials. The research on concept representation often needs to be based on a standardized and large-scale database of conceptual semantic features. Although Western scholars have established a variety of English conceptual semantic feature datasets, there is still a lack of a comprehensive Chinese version. In the present study, a Chinese Conceptual semantic Feature Dataset (CCFD) was established with 1,410 concepts including their semantic features and the similarity between concepts. The concepts were grouped into 28 subordinate categories and seven superior categories artificially. The results showed that concepts within the same category were closer to each other, while concepts between categories were farther apart. The CCFD proposed in this study can provide stimulation materials and data support for related research fields. All the data and supplementary materials can be found at https://osf.io/ug5dt/.

Keywords concept · semantic feature · dataset · Chinese

Introduction

When we think of "cat", we think of it as an animal, with a tail, fur, and being a pet, etc. These are our brain's memory storage and representation of the concept of "cat" and are expressed through the form of language, which are called the semantic features of a concept. Memory and language are very important high-level cognitive functions of humans. Researchers in many fields, such as psychologists, neuroscientists, and linguists, have been working hard to study their internal mechanisms. According to the conceptual structure account (Taylor, Devereux, & Tyler, 2011; Tyler & Moss, 2001), concepts can be expressed through their semantic features (e.g., "is hairy", "is an animal") and statistical measures, which are called conceptual structure statistics. Conceptual structure statistics can provide information about the parent category of a concept (for example, a cat is an animal) and how different a concept is within that category (for example, the uniqueness of the camel lies in its hump, which other animals do not have).

Concepts and their semantic features are important contents of memory. It has been found that distributed representations of knowledge about each attribute of the concept are in the corresponding areas of the brain (Binder, Desai, Graves, & Conant, 2009; Fernandino et al., 2016; Lambon Ralph, Jefferies, Patterson, & Rogers, 2017; Pulvermuller, 2013). For instance, the visual attributes (such as shape and color) are represented in ventral visual processing pathway including the occipital cortex, the ventral occipito-temporal cortex, the lateral occipital complex and fusiform gyri. The sound attributes (such as the sound made by an object) are associated with activation of the sound processing areas including the superior temporal gyrus and left ventrolateral

✉ Yaling Deng
yalingdeng@cuc.edu.cn

✉ Lihong Cao
lihong.cao@cuc.edu.cn

[1] State Key Laboratory of Media Convergence and Communication, Communication University of China, No.1 of Dingfuzhuang East Street, Chaoyang District, Beijing, China

[2] Neuroscience and Intelligent Media Institute, Communication University of China, Beijing 100024, China

[3] TV School, Communication University of China, Beijing 100024, China

[4] Animation and Digital Arts school, Communication University of China, Beijing 100024, China

[5] College of Humanities, Communication University of China, Beijing 100024, China

[6] State Key Laboratory of Mathematical Engineering and Advanced Computing, Wuxi 214125, China

prefrontal cortex (vlPFC). Whether the concept can be manipulated is represented in the areas processing motion including junction of the left posterior medial temporal gyrus (pMTG) and anterior occipital cortex (Fernandino et al., 2016; Lambon Ralph et al., 2017). Moreover, the human brain has two conceptual representation systems based on sensation and language system, respectively (Wang, Men, Gao, Caramazza, & Bi, 2020). When people recall information, they also search through the semantic network. Therefore, words that are semantically connected more closely with other words (which share greater semantic similarity with other words) are easier to be remembered (Xie, Bainbridge, Inati, Baker, & Zaghloul, 2020). However, it is still not clear how the human brain represents the relationship between concepts.

There are several computationally generating approaches to construct semantic spaces to represent the relationship between concepts. The word co-occurrence approach extrapolates words' relationship using their frequency of occurrence in the text (Van Rensbergen, De Deyne, & Storms, 2016). The word2vec algorithm uses a neural network model to learn word associations from a large corpus of text using vector with a particular list of numbers to represent each distinct word, so that mathematical function (such as cosine similarity) can be used to calculate the semantic similarity between words (Mikolov, Chen, Corrado, & Dean, 2013). Besides, several semantic databases have been established. WordNet is a lexical database of semantic relations between words in more than 200 languages, which links words into semantic relations including synonyms, hyponyms, and meronyms (George, 1995). HowNet is an online common-sense knowledgebase which puts more emphasis on the relationships between concepts (including inter-conceptual relationships and inter-attribute relationships) than WordNet (Dong, Dong, & Hao, 2006). These semantic networks constructed by using Internet big data have made outstanding contributions to many application fields of artificial intelligence, such as natural language processing (Skelac & Jandrić, 2020), information retrieval (Leydesdorff & Vaughan, 2006), intelligent question and answering (Yilmaz & Toklu, 2020), etc., and also provide rich data support and data modeling for research on how the human brain represents the relationship between knowledge (Armeni, Willems, & Frank, 2017).

However, in the research fields that investigate the neural mechanism of knowledge representation with humans as research objects, more researchers choose the data evaluated directly by humans as experimental materials (De Deyne, Navarro, Perfors, Brysbaert, & Storms, 2019; Jouravlev & Mcrae, 2016; Scott, Keitel, Becirspahic, Yao, & Sereno, 2019). Though these approaches are more time-consuming than the computationally generating norms mentioned above, the semantic space reflected by these materials may be the closest to the representation of knowledge in the human brain, so it is more widely used in the fields of psychology and neuroscience. Two methods are generally used to investigate semantic representation in humans. One way is to give a word as a clue, and let the subjects associate some other words freely to generate the relationship between the concepts (De Deyne et al., 2019), such as English Small World of Words project (SWOW-EN) (See https://smallworldofwords.org). This word association approach reveals mental representations that cannot be reduced to lexical usage patterns. Another way to measure the relationship between concepts is mainly based on the features shared between the concepts. The more features shared between concepts, the higher the similarity is (Cree & McRae, 2003; Griffiths, Steyvers, & Tenenbaum, 2007; Jones & Mewhort, 2007; Vigliocco, Vinson, Lewis, & Garrett, 2004). Generally, studies on semantic memory ask subjects to list the features of concepts, called feature production norms (Toglia, 2009), and then investigate how these concepts are represented in memory according to the features of the concepts listed by the subjects. Although a lot of related research has been done, most researchers collected the concepts and features individually instead of using a normed dataset (Ashcraft, 1978; Collins & Loftus, 1988; Mcnorgan, Kotack, Meehan, & Mcrae, 2007; Mcwilliams & Schmitter-Edgecombe, 2008; Toglia, 2009). Due to the difference of selected concepts and subjects, this self-normed research cannot be compared between each other directly. Therefore, the availability of stimulating materials greatly limits the related research.

The study of human brain knowledge representation needs to be based on a standardized and large-scale conceptual semantic feature database. At present, several English conceptual semantic feature datasets have been established. The earliest dataset was established by McRae, Cree, Seidenberg, and McNorgan (2005), who asked participants to describe 541 nouns, including animate objects (such as crocodiles) and inanimate objects (such as airplanes). Subsequently, Vinson and Vigliocco (2008) added verbs to the dataset, and collected 456 semantic features of nouns and verbs. Buchanan, Holmes, Teasley, and Hutchison (2013) collected 1,808 nouns, verbs, and adjectives. Recently, after years of hard work, Buchanan, Valentine, and Maxwell (2019) have released the latest and largest conceptual semantic feature dataset to date, containing 4,436 concepts, including nouns, verbs, adjectives, and other types of vocabulary. The above-mentioned English conceptual semantic feature datasets adopted a free description method, that is, given a word, the subjects could freely describe the features of the concept without any hints, and they wrote down whatever features they thought of.

Devereux, Tyler, Geertzen, and Randall (2014) added some concepts on the basis of the dataset established by McRae et al. (2005), and expanded the dataset to 638 concepts. Instead of letting the subjects describe freely, they adopted the way of guiding word prompt. The guiding words were: *is, has, does, made of*, etc. The subjects could fill in the corresponding features after each guiding word. According to their research results, they concluded that the data collection method with guiding words yielded more information than the free description method, and they designed a web system that could expand the dataset continuously.

Although there are not many English conceptual semantic feature datasets, the number of Chinese datasets is more limited. Currently, only two Chinese datasets have been established. One was assessed by Xiang, Lin, and Jiang (2015), which contains only 50 nouns. The other one, assessed by Gao, Lin, Jiang, and Lu (2016), contains only 30 verbs. Both datasets used the free description collection method. Compared with the English datasets, the related work of Chinese datasets is far behind.

For language research, there are great differences in different languages. At the same time, in the study of memory, people from different cultural backgrounds also have great differences in concept coding. Firstly, different categories of concepts have different typical representatives. For example, when it comes to the concept of the word "bird", Chinese are more familiar with the sparrow, while North Americans may first think of the robin. Secondly, the familiarity of different concepts is different for people with different cultural backgrounds. For example, the concept of "food" is a familiar superior category to everyone, but its subordinate categories (various foods) vary greatly among different cultures. People from different countries are only familiar with their own food ("粽子-zongzi" is familiar to Chinese but not to American). Finally, for the same concept, people from different cultural backgrounds may encode it differently, such as "the cross (十字架)". Therefore, the study of the Chinese language system and the study of the language, memory, and other cognitive functions of native Chinese speakers need to use the Chinese conceptual semantic feature dataset as the stimulus material, rather than the English ones. Due to the limitation of the current Chinese database, it is necessary to establish a Chinese conceptual semantic feature dataset for relevant researchers to promote the research on language and memory of Chinese, which is the main purpose of this study.

The present study collected semantic features of 1,410 concepts in Chinese by using guiding words. We provided the features (including the frequency of each feature) of each concept and calculated the similarities between each two concepts. Concepts were further grouped into different categories and compared between categories. Hierarchical clustering was then used to examine the relationship between concepts. We also compared this dataset with a widely used English version of conceptual semantic feature dataset (Devereux et al., 2014) to see whether the present data make sense and the difference between the Chinese and English versions.

Methods

Participants

Two hundred and four subjects (44 males) participated in this experiment, aged between 18 and 57 (M = 23.495, SD = 4.806). All the participants were Chinese, and their native language was Chinese. We also collected the region of each participant, and the participants came from 25 provinces (116 from the north of China, 88 from the south of China, see Fig. S1 in the "Supplementary materials_1_Figures" for detailed geographical distribution of participants). On average, each participant completed the assessment of 202 concepts. The participants were paid after completing the tasks.

Materials

A total of 1,410 concepts were selected as experimental materials in this study. In order to cover common concrete concepts as comprehensively as possible, we integrated multiple word datasets: 1) concepts in the dataset created by Devereux et al. (2014) (hereinafter referred to as CSLB); 2) concepts in ImageNet (Deng, Dong, Socher, Li, & Li, 2009); 3) concepts in the open Chinese lexicon (THUOCL) (Han et al., 2016) with word frequency exceeding 10,000 in IT, finance and economics, place names, historical celebrities, medicine, diet, law, automobile and animal category. The English parts of these concepts (CSLB and ImageNet) were translated into Chinese, and the duplicated concepts were deleted. Finally, we got 11,160 concepts. Three graduate students (all are native Chinese speakers) were then asked to rate the familiarity (from 1 to 9, 1 for complete unfamiliarity and 9 for very familiar) and concreteness (from 1 to 9) of each concept. The concreteness was defined as whether the concept was perceptible (for example, it can be seen or touched) and easy to be imagined (the more concrete a concept is, the easier it is to be imagined), so 1 means very abstract, imperceptible, and hard to be imagined, while 9 means very specific, perceptive, and easy to be imagined. The rating results of familiarity and concreteness of each concept are shown in the "Supplementary materials_2_Table1". We selected 1,410 concepts with the highest familiarity (M = 8.85, SD = 0.53) and concreteness (M = 8.97, SD = 0.22) scores as the formal experimental materials.

Procedure

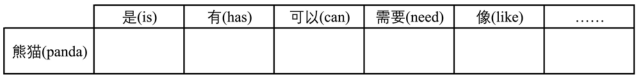
The research work of Devereux et al. has shown that the guiding words method can collect more comprehensive data than the free description way (Devereux et al., 2014). Before we formally collected data, we did a pre-experiment and compared the way with guiding words and the way of free description. Two groups of participants (*N* = 40) were recruited, and one group (n = 20) used the way of guiding word, while the other group (n = 20) used the way of free description. The instruction of guiding word way was shown below. The only difference between the two methods was that no guiding words were used in the free description group. We tested 15 concepts. The results showed that the guiding word group spent more time (15 min) than free description group (12 min). Besides, the guiding word group provided more features per concept (n = 25.87) than free description group (n = 17.00). Thus, the guiding word way may give participants more clue to search features from memory. Therefore, guiding word method was used in the present study.

All data collection work was carried out through the Internet. First, the participants were presented with experimental task introduction in simplified Chinese. The instruction is as follows:

> *Hello, thanks for your participation. You will see some words, and please describe each word, such as its appearance, sound, smell, touch, or how to use it, when and where to use it. You can also describe where it comes from and the feelings it brings to you. Please fill in the corresponding content under each guiding word (is-*是*, has-*有*, can-*可以*, need-*需要*, like-*像*, ...-*其他*), and you can fill in any content you want in the box below "......" (ellipsis).*
>
> *Note:*

1. *Please fill according to your intuition;*
2. *Multiple words can be filled under each guiding word, separated by spaces;*
3. *Please give all the features or attributes you can think of as comprehensively as possible;*
4. *If you don't know what this word means, or you don't know what feature it has, you can leave it blank.*

When the participants understood the experimental task, the formal experiment started. For each concept, the interface is shown in Fig. 1. Participants need to fill in the blanks with the features of each concept (without time limit). Referring to the data collection method of CSLB (Devereux et al., 2014), this study selected six guiding words: "is-是, has-有, can-可以, need-需要, like-像, ...-其他". Participants can fill in the corresponding features under the corresponding guiding words. "……" means "other", which can be filled in any features below. The concepts were presented completely random among the participants. Each concept was evaluated by at least 30 participants.



	是(is)	有(has)	可以(can)	需要(need)	像(like)	……
熊猫(panda)						

Fig. 1 Schematic diagram of data collection

Data analysis

Preprocessing

For all the 1,410 concepts, we got a total of 381,119 original descriptions. Each of the original description was checked one by one and preprocessed (if necessary) as follows:

(1) Correct typos. Because the data were collected electronically, most participants used pinyin input method, which led to typos, such as "香蕉是职务 (banana is a job)", but actually is "香蕉是植物 (banana is a plant)". (In Chinese, "job-职务" and "plant-植物" are both pronounced "zhiwu".)

(2) Delete completely wrong descriptions (those are completely against the common sense), punctuation marks, strange characters, etc. For example, "电灯是食物 (electric light is food)".

(3) Match the feature description to the correct guiding word. During the data collection, there were multiple guiding words, and some participants did not fill into the guiding words properly. For example, "猫像动物 (cats are like animals)" was changed to "猫是动物 (cats are animals)".

(4) Delete adverbs of degree, such as "very". For example, "热狗是非常香的 (hot dogs are very fragrant)" was changed to "热狗是香的 (hot dogs are fragrant)".

(5) Unify expressions. In order to facilitate subsequent analysis, we referred to the data preprocessing methods of CSLB (Devereux et al., 2014), and unified the expressions of the inspected data. Many features were described in different words, but the actual meaning was the same (synonyms). The features with the same meaning were unified into the same statement. For instance, "是食物 (is food)", "是食品 (is food)", "是吃的 (is for eaten)", "是食用的 (is edible)" were all unified into "是食物 (is food)".

(6) Feature segmentation. Split combined features into independent features referred to the method used in CSLB (Devereux et al., 2014). Some participants used the form of "adjective + noun" to describe concepts, such as: "玫瑰是红色的花 (rose is a red flower)", which contains two features, which can be split into "玫瑰是红色的 (rose is red)" and "玫瑰是花 (rose is a flower)". Another example is "汽车有四个轮子 (car has four wheels)", which can be split into "汽车有轮子 (car has wheels)" and "汽车有四个轮子 (car has four wheels)".

After preprocessing, there were 378,533 descriptions left (2,586 fewer than original descriptions, which is 0.006%). All the following processing and related results were based on the data after preprocessing.

Within concepts

After the preprocessing, we calculated the overall information of the Chinese Conceptual semantic Feature Dataset (CCFD), including the number of effective feature descriptions and the average number of features per concept. Then we counted the features of each concept, how many participants effectively described each concept (after deleting invalid descriptions), and the frequency of each feature of each concept (that is, how many participants mentioned this feature). In addition, the sample coverage of each concept was calculated (Canessa, Chaigneau, Lagos, & Medina, 2020). The higher the sample coverage is, the more completeness the features are. Since the number of participants describing each concept after data preprocessing was not the same, we normalized the frequency information of each concept by "(frequency of each feature of the concept / number of participants describing the concept) × 30".

Between concepts

In this study, we would like to investigate the features common to multiple concepts (shared features) and the features unique to certain concepts (unique features). We calculated the sharing degree of each feature, that is, the number of concepts that have this feature. The features only belong to one or two certain concepts were defined as unique features. Features possessed by three or more concepts were shared features, and the classification criteria were the same as CSLB (Devereux et al., 2014). Then, we investigated the similarity between concepts, which was a key point of this study. The 1,410 concepts have a total of 10,059 features; thus, each concept has a 10,059-dimension feature vector with each element corresponding to the number of participants describing this feature. Principle component analysis (PCA) was used to centralize and reduce the dimension of feature vector to 645, which explained 95% of the variance, then the cosine distance was used to calculate the similarity between concepts.

Between categories

We investigated whether the concepts of different categories were different within and between categories. There is no uniform standard for the classification of concepts. Some studies on patients with brain injury have found that there are obstacles to the identification of specific categories of concepts, namely selective semantic impairment (Gainotti, 2000, 2005; Gainotti, Spinelli, Scaricamazza, & Marra, 2013). For example, the vast majority of patients with semantic disorders, who have lesions in the mid-to-anterior ventral and medial temporal regions bilaterally, are mainly restricted to the animal domain but not to plants or artificial objects (Gainotti, 2010), while damage to the left medial and inferior temporal regions hinders the recognition of plants (Samson & Pillon, 2003). In addition to these categories of living things (animals, plants) and non-living things (artificial objects, natural objects), body parts are usually treated as a separate category (Guido, 2015). According to these selective semantic impairment studies, we identified five superior categories: animal, artificial object, body part, plant, and natural object. During the actual processing of classification, we found that the food did not belong to any of the above categories (though it could be sorted into artificial object, but many plants also belong to food), so we made it a separate category. Although transportation could be directly classified as artificial object, it could also be further divided into vehicles, boats, and aircrafts. Therefore, for the convenience of subsequent classification, we also took transportation as a superior category. As a result, all concepts were artificially divided into seven superior categories as: animal, transportation, artificial object, body part, food, plant, and natural objects. To further distinguish more detailed conceptual categories, we artificially divided the seven superior categories into 28 subcategories (see Table 4 for details. “Supplementary materials_2_Table1” shows the category of each concept). Both the classification of superior categories and subcategories were determined by discussion among the authors.

Moreover, the ward minimum variance algorithm was used for hierarchical clustering. Unlike the artificially categorization approach according to domains, this is a data-driven approach based on the features of concepts. We compared the results of the two classifications.

Studies have shown that different categories of objects have different numbers of shared features. For example, animals have more shared features (such as having eyes, ears, and nose), while tools have fewer shared features but more unique features (Clarke & Tyler, 2015). Human use different brain regions to recognize different objects. Recognizing objects with more shared features requires more participation of the lateral posterior fusiform gyrus, while recognizing objects with fewer shared features relies more on the medial posterior fusiform gyrus (Tyler et al., 2013). Concepts that share more features and prone to confusion are difficult to distinguish, such as 海象 (walrus) and 海狮 (sea lion), which require more involvement of the entorhinal cortex (Bruffaerts et al., 2013). We compared the number of total features, number (and proportion) of shared features, and number (and proportion) of unique features between different subcategories and superior categories. Because the subcategories were nested into the superior categories, so we used the general linear model to analyze the nested data.

Results

Within concepts

In CCFD, there are 378,533 valid descriptions for 1,410 concepts, with an average of 268 descriptions for each concept, and an average of nine descriptions for each concept given by each participant. The sample coverage of each concept is presented in the "Supplementary material_2_Table1" and the average sample coverage is 68.84%. The features of each concept and the frequency of each feature are presented in the "Supplementary material_3_Table2_All concepts and features". Among them, the features with frequency equaling to 1 were deleted as these features are not universal, which is the same as the processing method of CSLB (Devereux et al., 2014). Table 1 shows a schematic result of the concept "熊猫 (panda)". The average number of features of each concept which belongs to CCFD and CSLB respectively are shown in Table 2.

Between concepts

In this study, the concepts and features generated by normalized results were used to form a concept feature matrix (see the "Supplementary material_4_Table3_Concept feature matrix" for details). The 1,410 concepts have a total of 10,059 features. On average, each feature is shared by 5.17 (SD = 22.71) concepts. Table 2 shows the average number of shared features and unique features for each concept of both CCFD and CSLB. We divided the number of shared features according to the sharing degree, as shown in Table 3.

The similarity results between concepts are presented in the "Supplementary material_5_Table4_Concept similarity matrix". The concepts used in this study have partially overlap with CSLB. The overlapping concepts cover a total of 19 subcategories. We randomly selected six concepts from each subcategory to calculate the similarity between concepts, as shown in Fig. 2. Figure 2a shows the result of CCFD, and Fig. 2b shows the result of CSLB. We calculated the similarities within and between each superior category of both CCFD and CSLB (see Fig. S2 in "Supplementary materials_1_Figures" for details). The results showed that, compared with CSLB, CCFD had higher average intra-category similarity (CCFD: 0.26 > CSLB: 0.23) and lower average inter-category similarity (CCFD: – 0.04 < CSLB: 0.03). Cosine distance was used to calculate the correlation between CCFD and CSLB, and the correlation between the two similarity matrices based on off-diagonal pairs was 58.29% (including all the shared concepts).

Between categories

This study also hierarchically clustered all concepts based on the features of the concepts. The results of hierarchical clustering are detailed in the "Supplementary materials_6_Table5_Results of hierarchical clustering". For the whole view of hierarchical clustering results, please see "Supplementary materials_7_Results of hierarchical clustering". All concepts were grouped into two large clusters, which can be named as living things and non-living things, respectively. Living things were further divided into two clusters: animals and plants. We compared the clustering results and human category judgments. The results showed that the "animal" and "plant" categories were the same between the two methods of classification. However, the concepts which

Table 1 The features and the corresponding frequency of the concept "panda" as an example

Features	Frequency	Features	Frequency	Features	Frequency
need-bamboo	23	like-selenarctos thebetanus	7	is-toy	2
is-animal	22	need-air	7	like-cat	2
need-food	21	need-sleep	7	like-plush toys	2
is-National treasure	18	has-ears	7	like-lesser panda	2
like-bear	16	has-mouth	7	need-care	2
need-water	16	is-mammal	6	need-breath	2
can-climb trees	14	is-cute	6	need-fed	2
is-black	12	has-limbs	6	need-sunlight	2
is-white	11	can-acting cute	4	has-brain	2
need-protection	11	has-nose	4	has-many kinds	2
has-eyes	11	can-produce	3	has-feet	2
is-black and white	10	is-felidae	3	has-hand	2
has-black eye	8	has-legs	3	has-tail	2
has-hair	8	has-claws	3	has-teeth	2
can-be viewed	7	other-charmingly naive	2		

Table 2 Dataset overview of CCFD and CSLB (Devereux et al., 2014)

	Number of concepts	Number of features per concept	Number of shared features (≥ 3)	Number of unique features (≤ 2)
CCFD	1,410	36.85 ± 7.63	30.67 ± 7.93	6.18 ± 3.86
CSLB	638	35.53 ± 7.20	27.12 ± 6.52	8.41 ± 4.86

belong to "transportation, artificial object, body part, food and natural object" judged by human were put into a big cluster in the hierarchical clustering results. Interestingly, we could find some associations in the results of hierarchical clustering. For example, many kinds of vehicles (cars, ships, planes) and many kinds of roads and bridges gathered in a cluster, which are related to transportation. The various types of pens and hand-related concepts (hand, arm, finger, palm) also clustered to reflect the association of using these pens by hands.

We also compared the hierarchical clustering results between CCFD and CSLB. Because there were too many concepts, it was impossible to visualize them all here. Therefore, we only visualized the hierarchical clustering results of the animal category (Fig. 3). The results showed that CCFD had a similar hierarchical structure in the animal category as CSLB, which could be well divided into 鸟 (bird) (green lines and words), 昆虫 (insect) (red lines and words), 陆生动物 (terrestrial animal) (blue lines and words) and 水生动物 (aquatic animal) (purple lines and words). As we can see from Fig. 3b, some animals that live in the water (such as shark and dolphin) were clustered with terrestrial animals but not with aquatic animals in the CSLB.

The number of concepts in each category, the number of shared features, unique features, and total features of the concepts were analyzed in 28 subcategories and seven superior categories, as shown in Table 4.

Considering the nested relationship between subcategory and superior category, nested analysis of UNIANOVA was used to examine the difference in the number of total features, number (and proportion) of shared features and unique features of the concepts of different subcategories as well as superior categories. It was found that there were significant differences in the number of shared features ($F_{(21,1382)}$ = 7.173, $p < 0.001$, $\eta^2 = 0.098$), proportion of shared features ($F_{(21,1382)} = 6.572$, $p < 0.001$, $\eta^2 = 0.091$), number of unique features ($F_{(21,1382)} = 4.253$, $p < 0.001$, $\eta^2 = 0.061$), proportion of unique features ($F_{(21,1382)} = 6.572$, $p < 0.001$, $\eta^2 = 0.091$) and the total number of features ($F_{(21,1382)} = 3.956$, $p < 0.001$, $\eta^2 = 0.057$) among subcategories. The results also showed significant differences in the number of shared features ($F_{(6,1382)} = 72.330$, $p < 0.001$, $\eta^2 = 0.239$), proportion of shared features ($F_{(6,1382)} = 81.682$, $p < 0.001$, $\eta^2 = 0.262$), number of unique features ($F_{(6,1382)} = 51.622$, $p < 0.001$, $\eta^2 = 0.183$), proportion of unique features ($F_{(6,1382)} = 81.682$, $p < 0.001$, $\eta^2 = 0.262$) and the total number of features ($F_{(6,1382)}$ = 27.817, $p < 0.001$, $\eta^2 = 0.108$) among superior categories. Figure 4 shows the proportion of shared features and unique features of each subcategory (a) and superior category (b).

Because there are too many subcategories in this study (n = 28), we did not report the post hoc LSD multiple comparison results of subcategory but only reported the results of superior category. For the total number of features, animals ≥ transportation ≥ plants > food ≥ natural objects ≥ artificial objects ≥ body parts (">" means significant difference, $p < 0.05$, "≥" means no significant difference). There was no significant difference between animals and transportation ($p = 0.078$), transportation and plants ($p = 0.113$), food and natural objects ($p = 0.658$), food and artificial objects ($p = 0.055$), natural objects and artificial objects ($p = 0.677$), artificial objects and body parts ($p = 0.481$), body parts and food ($p = 0.115$), body parts and natural objects ($p = 0.421$), but there were significant differences between the other pairs.

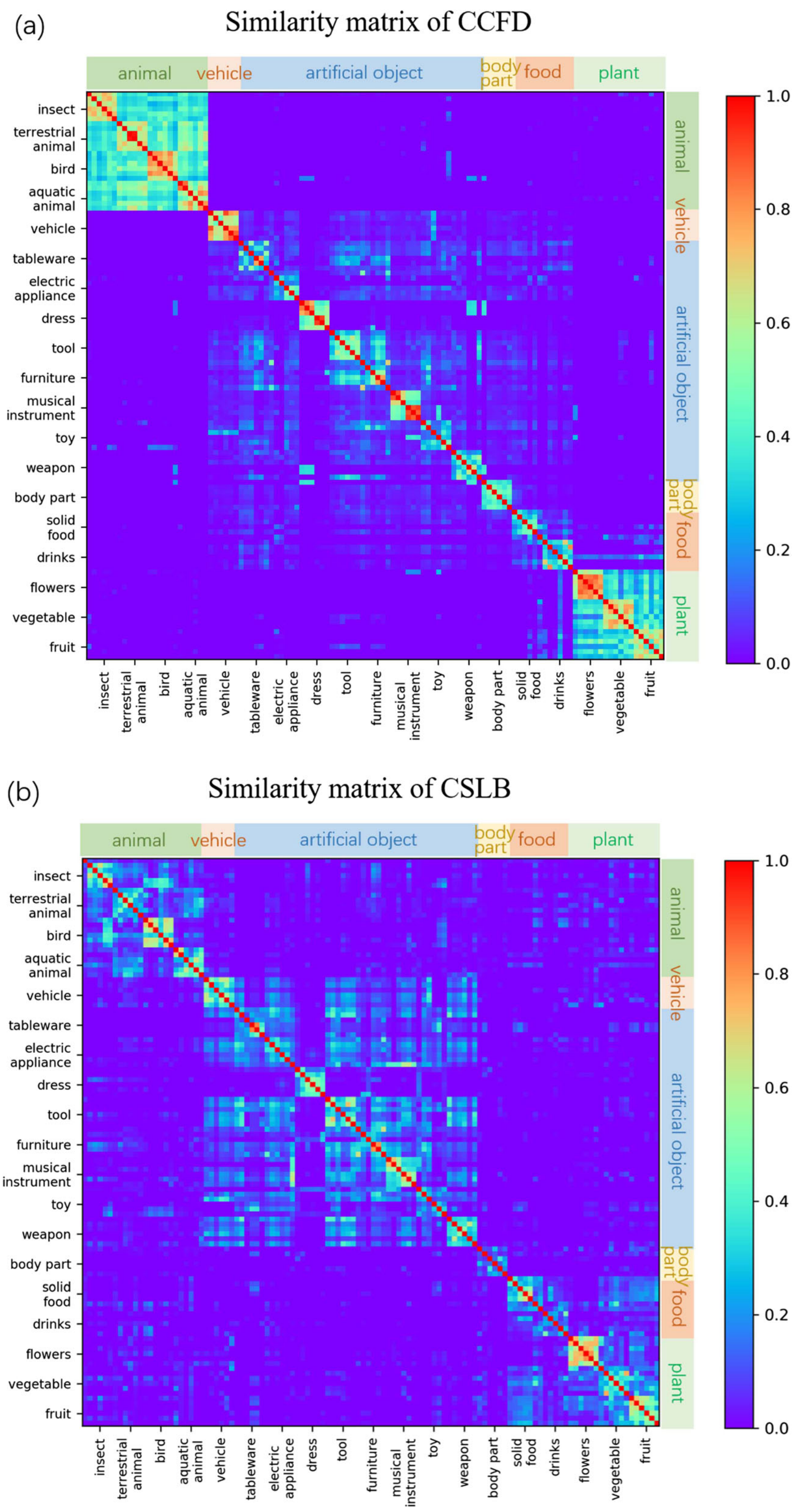
For the proportion of shared features, plants ≥ animals (p = 0.426) > food > transportation > artificial objects > natural objects ≥ body parts ($p = 0.143$). The results of unique feature proportion were just the opposite of shared features. From the data we could see that animals and plants had relatively the most shared features and the fewest unique features, while natural objects and body parts had relatively the most unique features and the fewest shared features.

Discussion

This study established the first Chinese conceptual semantic feature dataset with 1,410 concepts. All the concepts are very familiar to Chinese native speakers. The CCFD covers seven superior categories and 28 subcategories, and each concept has about 37 features. According to the features of concepts, the similarity analysis results showed

Table 3 The number of features with different sharing degree

Sharing degree	> 200	150~199	100~149	50~99	10~49	3~9	2	1
Number of features	24	21	26	82	651	1,911	1,370	5,974



(a) Similarity matrix of CCFD

(b) Similarity matrix of CSLB

Fig. 2 Similarity matrix of randomly selected concepts among 19 subcategories in CCFD (**a**) and CSLB (**b**). Each cell in the picture represents the similarity between two concepts. The concepts from the same subcategory (six randomly selected concepts from each subcategory) gathered together, and the text on the left and bottom of the picture represents the name of the subcategory (*black words*). The concepts from the same superior category also gathered together. The *upper* and *right sides* of the picture represent the names of the superior categories (*colored words*)

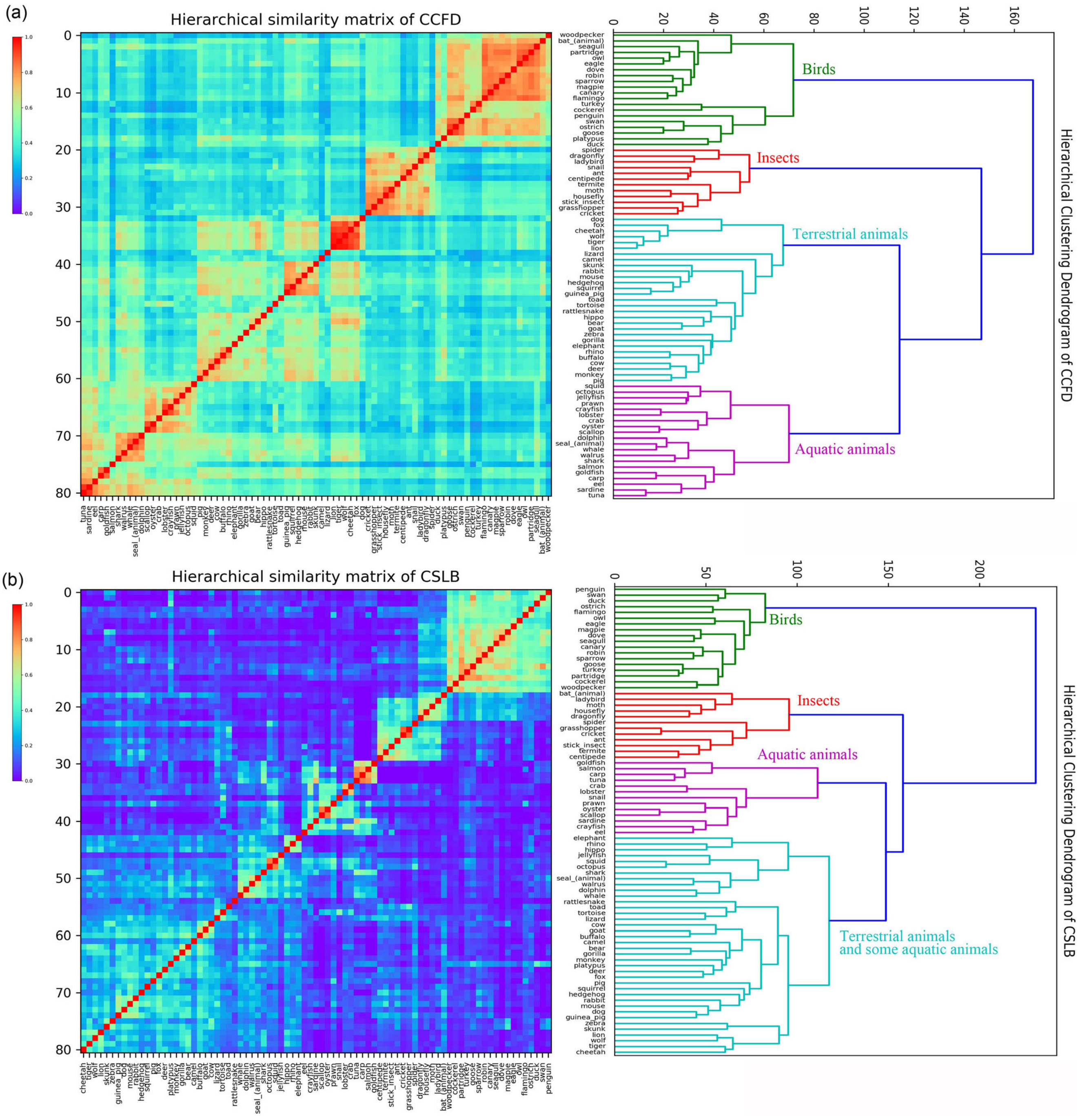


Fig. 3 The results of hierarchical clustering of animals. (a) The results of CCFD. (b) The results of CSLB. The concepts chosen by the two are the same.

that the concepts in the same category were similar, while the concepts between different categories were less similar, but they were not completely irrelevant, which is consistent with our daily understanding of objects. Even for objects from different categories, we can still generate certain associations in non-category relationships, such as chopsticks and noodles that often appear in the same scene, the former belongs to tableware while the latter belongs to food.

Although there are several English conceptual semantic feature datasets (Buchanan et al., 2013; Buchanan et al., 2019; Devereux et al., 2014; McRae et al., 2005; Vinson & Vigliocco, 2008), the CCFD proposed in this study is the first large-scale conceptual semantic feature dataset in Chinese. Compared with the existing standardized English dataset, the CCFD is considerable both in terms of data quantity and quality. The biggest difference between CCFD and English dataset lies in the specific concepts and features, especially in the

Table 4 The number of concepts, shared features, unique features, and total features of the concepts in 28 subcategories and seven superior categories

Superior category	Subcategory	Number of concepts	Number of shared features	Number of unique features	Number of features
animal	insect	32	34.94 ± 7.69	4.91 ± 3.80	39.84 ± 8.19
	terrestrial animal	78	40.38 ± 6.56	4.54 ± 2.72	44.92 ± 7.21
	bird	33	36.61 ± 5.88	4.12 ± 2.23	40.73 ± 6.30
	aquatic animal	53	37.77 ± 6.60	3.77 ± 2.21	41.55 ± 6.71
	total	196	38.15 ± 6.92	4.32 ± 2.74	42.47 ± 7.35
transportation	vehicle	35	33.60 ± 6.49	6.86 ± 3.93	40.46 ± 7.29
	boat	7	33.57 ± 12.61	9.57 ± 3.36	43.14 ± 11.85
	aircraft	11	30.73 ± 4.45	8.36 ± 3.38	39.09 ± 5.61
	total	53	33.00 ± 7.14	7.53 ± 3.83	40.53 ± 7.64
artificial object	tableware	28	32.25 ± 7.24	6.04 ± 3.11	38.29 ± 7.40
	electric appliance	88	26.86 ± 6.56	7.59 ± 3.99	34.45 ± 8.51
	dress	99	30.59 ± 6.00	5.70 ± 3.00	36.28 ± 6.20
	tool	261	26.52 ± 6.43	8.27 ± 3.97	34.79 ± 7.01
	furniture	41	30.68 ± 7.13	6.10 ± 3.36	36.78 ± 7.57
	building	28	23.29 ± 5.75	9.11 ± 4.13	32.39 ± 5.18
	musical instrument	21	25.48 ± 6.15	5.76 ± 2.79	31.24 ± 6.45
	toy	31	23.77 ± 6.18	7.23 ± 4.39	31.00 ± 7.03
	weapon	26	29.54 ± 4.55	6.04 ± 2.42	35.58 ± 5.88
	medical supplies	15	25.13 ± 6.59	8.06 ± 3.24	33.20 ± 5.31
	sports equipment	18	26.61 ± 7.10	7.50 ± 3.65	34.11 ± 8.40
	total	656	27.48 ± 6.74	7.35 ± 3.83	34.83 ± 7.18
body part	body part	44	23.50 ± 7.14	10.55 ± 3.84	34.05 ± 6.75
food	solid food	100	31.87 ± 6.09	4.75 ± 2.85	36.62 ± 6.65
	drinks	58	30.47 ± 5.59	5.41 ± 3.12	35.88 ± 6.77
	meat	31	30.61 ± 6.23	4.29 ± 2.60	34.90 ± 7.17
	seasoning	25	28.84 ± 6.93	5.48 ± 2.45	34.32 ± 7.41
	total	214	30.95 ± 6.12	4.95 ± 2.86	35.90 ± 6.85
plant	flower	58	31.05 ± 4.80	3.83 ± 2.26	34.88 ± 5.42
	vegetable	71	37.25 ± 6.07	3.15 ± 2.44	40.41 ± 5.95
	tree	25	33.60 ± 6.59	5.00 ± 2.68	38.60 ± 6.30
	fruit	57	37.37 ± 5.79	3.47 ± 2.54	40.84 ± 6.20
	total	211	35.15 ± 6.34	3.64 ± 2.50	38.79 ± 6.39
natural object	natural object	36	25.50 ± 9.08	9.83 ± 4.77	35.33 ± 9.64

category of food. The category of food contains some concepts that people are very familiar with, but there are cultural differences for the familiarity of specific food, such as 臭豆腐 (stinky tofu), 皮蛋 (preserved eggs), 冰糖葫芦 (iced sugar gourd) and other concepts that only Chinese are familiar with. Therefore, CCFD is more suitable for research in Chinese. By comparing the Chinese and English conceptual feature datasets, researchers can also systematically investigate the differences between cultural differences, including the differences of concepts, the differences of features of the same concepts, etc.

By analyzing the frequency of each feature mentioned in a concept, we can see which features are core features and which features are marginal features. How to define a concept according to these features is also an important issue for language and cognitive psychologists. According to the features of these concepts, the concepts can be classified into different categories or even classified by hierarchy categories. However, how the human brain represents the category of concepts and the hierarchical relationship of concepts remains to be explored. The CCFD can provide data support for such studies. Our study also found that the number of shared features and unique features of different categories of concepts were different. Natural organisms such as animals and plants have more shared features, while the concepts of artificial things such as tools have fewer sharing features, which is consistent with previous research results (Clarke & Tyler, 2015).

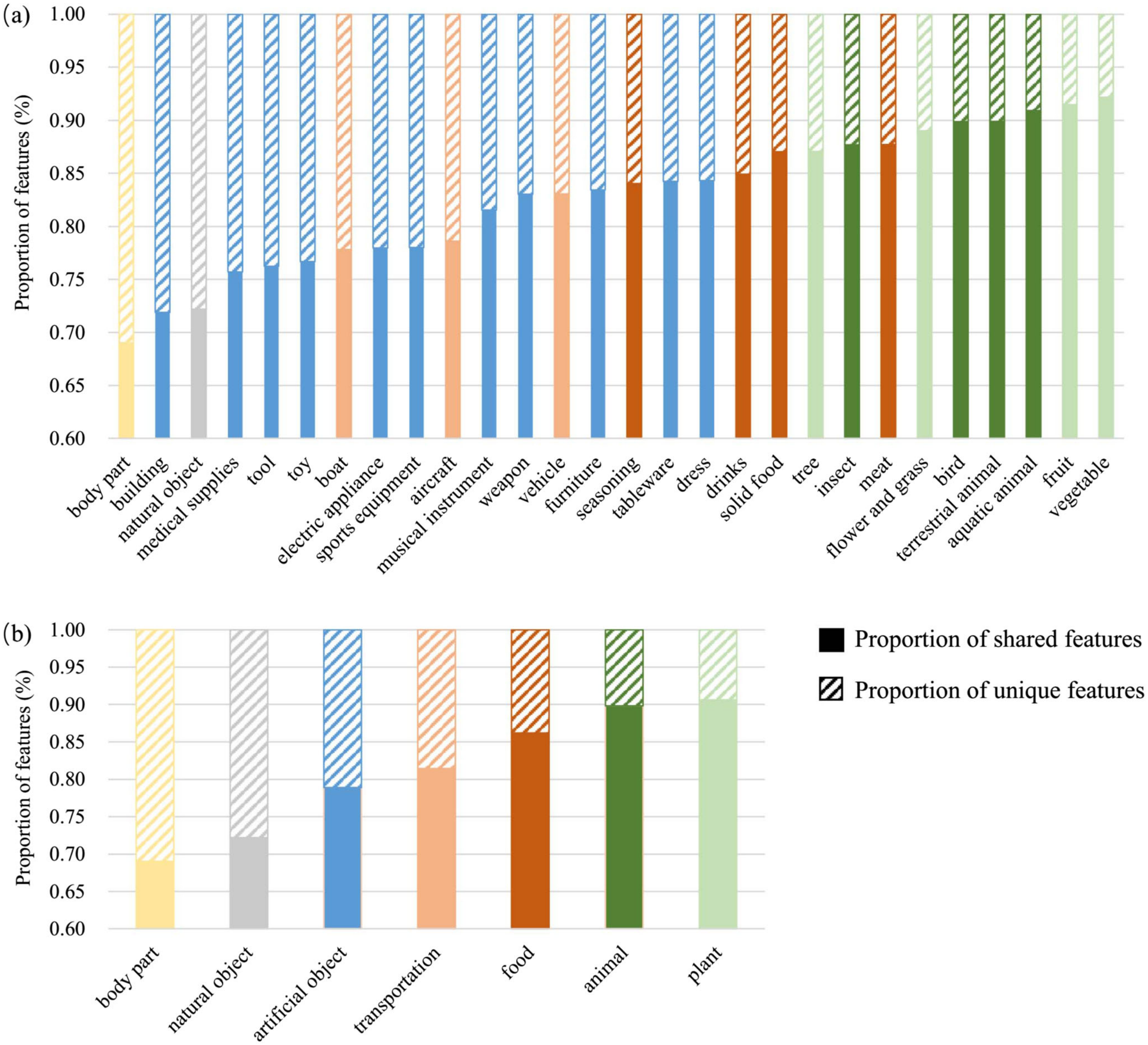


Fig. 4 Proportion of shared features and unique features of each subcategory (**a**) and superior category (**b**). The color in this figure is the same as the color index in Fig. 2

A lot of research on memory and language use semantic priming paradigm. Researchers manipulate priming effect by manipulating the correlation between priming words and target words. Compared with unrelated cues, related cues enable subjects to recognize target words more quickly (Schacter, Dobbins, & Schnyer, 2004). How to determine the degree of correlation between concepts, and then choose the appropriate priming words and target words, is the key to related research. In this study, the similarity between concepts was calculated by their features. Future research on semantic priming paradigm can use the present results as one of the indicators to measure the correlation between priming words and target words. In conclusion, the CCFD provides rich data support for semantic priming paradigm.

In addition to the standardization of datasets in language and memory research, computational modeling of memory also needs large-scale datasets to describe semantic memory accurately. The field of artificial intelligence also pays attention to the semantic features of concepts. Researchers try to use a knowledge graph as an abstract carrier to represent the cognitive content of knowledge on the dataset system, so as to serve the fields of machine translation and intelligent question and answering (Balaid, Abd Rozan, Hikmi, & Memon, 2016). The number of concepts in knowledge graph dataset is much larger than that in artificially evaluated dataset, but the features of these concepts are usually extracted from the massive Internet data, and the relationship between them is represented by logical structures. This is different from the human brain's learning and representation of concepts. The human brain does not acquire and store all the features of a concept simultaneously, but acquires each feature of the concept at different time in life. Moreover, the learning process can be divided into fast learning processes and slow learning processes. When the content of new learning information is consistent with the existing knowledge structure, it can be quickly integrated into the neocortex. When it is inconsistent with existing knowledge, it takes a long time to be processed in the medial temporal cortex and hippocampus, and then slowly integrated into the neocortex (Kumaran, Hassabis, & McClelland, 2016). Various modal features of the concept are represented distributely in the neocortex (Binder et al., 2009), such as the image of a dog in the visual cortex, the bark in the auditory cortex, and the semantic representation of the dog's features in the anterior temporal lobe (Wang et al., 2020). Although the representation of concepts by the human brain is not completely accurate, it has strong flexibility, which is also an important embodiment of human brain intelligence. Therefore, although there are massive datasets in the field of artificial intelligence and they can also construct a meaningful semantic space, they may not be suitable to study the representation of knowledge in the human brain. Instead, the representation of human brain knowledge may guide the establishment of a brain-like knowledge map to better serve the field of artificial intelligence.

We hope the CCFD can provide standardized data to support the researchers in the fields of linguistics, psychology, neuroscience, artificial intelligence, and so on. We will continue to expand the content of CCFD. On the one hand, we will expand the data collection for different age groups, such as the data for children of all ages, so as to investigate the learning process of concepts in the natural state; on the other hand, we will expand the number and types of concepts, including verbs and adjectives as well as abstract concepts, to further investigate the cognitive process of language.

Acknowledgements This study was supported by the Fundamental Research Funds for the Central Universities (Grant Nos. CUC200A004, CUC18A001, CUC18A003-3 and CUC2019B079), the High-Quality and Cutting-Edge Disciplines Construction Project for Universities in Beijing (Internet Information, Communication University of China), Ministry of Education of the PRC (20YJAZH085), and the Open Project Program of the State Key Laboratory of Mathematical Engineering and Advanced Computing (Grant No. 2020A09).

References

Armeni, K., Willems, R. M., & Frank, S. L. (2017). Probabilistic language models in cognitive neuroscience: Promises and pitfalls. *Neuroscience & Biobehavioral Reviews, 83*, 579–588. https://doi.org/10.1016/j.neubiorev.2017.09.001.

Ashcraft, M. H. (1978). Property norms for typical and atypical items from 17 categories: A description and discussion. *Memory & Cognition, 6*(3), 227–232. https://doi.org/10.3758/BF03197450

Balaid, A., Abd Rozan, M. Z., Hikmi, S. N., & Memon, J. (2016). Knowledge maps: A systematic literature review and directions for future research. *International Journal of Information Management, 36*(3), 451–475. https://doi.org/10.1016/j.ijinfomgt.2016.02.005

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex, 19*(12), 2767–2796. https://doi.org/10.1093/cercor/bhp055

Bruffaerts, R., Dupont, P., Peeters, R., De Deyne, S., Storms, G., & Vandenberghe, R. (2013). Similarity of fMRI activity patterns in left perirhinal cortex reflects semantic similarity between words. *The Journal of Neuroscience, 33*(47), 18597–18607. https://doi.org/10.1523/JNEUROSCI.1548-13.2013

Buchanan, E. M., Holmes, J. L., Teasley, M. L., & Hutchison, K. A. (2013). English semantic word-pair norms and a searchable Web portal for experimental stimulus creation. *Behavior Research Methods, 45*(3), 746–757. https://doi.org/10.3758/s13428-012-0284-z

Buchanan, E. M., Valentine, K. D., & Maxwell, N. P. (2019). English semantic feature production norms: An extended database of 4436 concepts. *Behavior Research Methods, 51*(4), 1849–1863. https://doi.org/10.3758/s13428-019-01243-z

Canessa, E., Chaigneau, S. E., Lagos, R., & Medina, F. A. (2020). How to carry out conceptual properties norming studies as parameter estimation studies: Lessons from ecology. *Behavior Research Methods*. https://doi.org/10.3758/s13428-020-01439-8

Clarke, A., & Tyler, L. K. (2015). Understanding What We See: How We Derive Meaning From Vision. *Trends in Cognitive Sciences, 19*(11), 677–687. https://doi.org/10.1016/j.tics.2015.08.008

Collins, A. M., & Loftus, E. F. (1988). A spreading-activation theory of semantic processing. *Psychological Review, 82*(6), 407–428. https://doi.org/10.1037//0033-295X.82.6.407

Cree, G. S., & McRae, K. (2003). Analyzing the factors underlying the structure and computation of the meaning of chipmunk, cherry, chisel, cheese, and cello (and many other such concrete nouns). *Journal of Experimental Psychology. General, 132*(2), 163–201. https://doi.org/10.1037/0096-3445.132.2.163

De Deyne, S., Navarro, D. J., Perfors, A., Brysbaert, M., & Storms, G. (2019). The "Small World of Words" English word association norms for over 12,000 cue words. *Behavior Research Methods, 51*(3), 987–1006. https://doi.org/10.3758/s13428-018-1115-7

Deng, J., Dong, W., Socher, R., Li, L. J., & Li, F. F. (2009). *Imagenet: A large-scale hierarchical image database.* Paper presented at the 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2009), 20–25 June 2009, Miami, Florida, USA.

Devereux, B. J., Tyler, L. K., Geertzen, J., & Randall, B. (2014). The Centre for Speech, Language and the Brain (CSLB) concept property norms. *Behavior Research Methods, 46*(4), 1119–1127. https://doi.org/10.3758/s13428-013-0420-4

Dong, Z., Dong, Q., & Hao, C. (2006). HowNet and the Computation of Meaning. Singapore: World Scientific. https://doi.org/10.1142/5935

Fernandino, L., Binder, J. R., Desai, R. H., Pendl, S. L., Humphries, C. J., Gross, W. L., … Seidenberg, M. S. (2016). Concept Representation Reflects Multimodal Abstraction: A Framework for Embodied Semantics. *Cerebral Cortex, 26*(5), 2018–2034. https://doi.org/10.1093/cercor/bhv020

Gainotti, G. (2000). What the locus of brain lesion tells us about the nature of the cognitive defect underlying category-specific disorders: a review. *Cortex, 36*(4), 539–559. https://doi.org/10.1016/S0010-9452(08)70537-9.

Gainotti, G. (2005). The influence of gender and lesion location on naming disorders for animals, plants and artefacts. *Neuropsychologia, 43*(11), 1633–1644. https://doi.org/10.1016/j.neuropsychologia.2005.01.016

Gainotti, G. (2010). The influence of anatomical locus of lesion and of gender-related familiarity factors in category-specific semantic disorders for animals, fruits and vegetables: a review of single-case studies. *Cortex, 46*(9), 1072–1087. https://doi.org/10.1016/j.cortex.2010.04.002.

Gainotti, G., Spinelli, P., Scaricamazza, E., & Marra, C. (2013). The evaluation of sources of knowledge underlying different semantic categories. *Frontiers in Human Neuroscience, 7*, 40. https://doi.org/10.3389/fnhum.2013.00040

Gao, J., Lin, F., Jiang, Z., & Lu, H. (2016). The modeling and analysis of semantic features for Chinese verbs. *Chinese Journal of Rehabilitation Medicine, 31*(4), 381–387.

George, A. M. (1995). WordNet: A Lexical Database for English. *Communications of the ACM, 38*(11), 39–41. https://doi.org/10.1145/219717.219748.

Griffiths, T. L., Steyvers, M., & Tenenbaum, J. B. (2007). Topics in semantic representation. *Psychological Review, 114*(2), 211–244. https://doi.org/10.1037/0033-295X.114.2.211.

Guido, G. (2015). Inborn and experience-dependent models of categorical brain organization. A position paper. *Frontiers in Human Neuroscience, 9*, 2. https://doi.org/10.3389/fnhum.2015.00002

Han, S., Zhang, Y., Ma, Y., Tu, C., Guo, Z., Liu, Z., Sun, M. (2016). THUOCL: Tsinghua Open Chinese Lexicon. Tsinghua University.

Jones, M., & Mewhort, D. (2007). Representing word meaning and order information in a composite holographic lexicon. *Psychological Review, 114*(1), 1–37. https://doi.org/10.1037/0033-295X.114.1.1.

Jouravlev, O., & Mcrae, K. (2016). Thematic relatedness production norms for 100 object concepts. *Behavior Research and Methods*, 48, 1349–1357. https://doi.org/10.3758/s13428-015-0679-8.

Kumaran, D., Hassabis, D., & McClelland, J. L. (2016). What Learning Systems do Intelligent Agents Need? Complementary Learning Systems Theory Updated. *Trends in Cognitive Sciences, 20*(7), 512–534. https://doi.org/10.1016/j.tics.2016.05.004

Lambon Ralph, M. A., Jefferies, E., Patterson, K., & Rogers, T. T. (2017). The neural and computational bases of semantic cognition. *Nature Reviews. Neuroscience, 18*(1), 42–55. https://doi.org/10.1038/nrn.2016.150

Leydesdorff, L., & Vaughan, L. (2006). Co-occurrence matrices and their applications in information science: Extending ACA to the Web environment. *Journal of the American Society for Information Science and Technology, 57*(12), 1616–1628. https://doi.org/10.1002/asi.20335

Mcnorgan, C., Kotack, R. A., Meehan, D. C., & Mcrae, K. (2007). Feature-feature causal relations and statistical co-occurrences in object concepts. *Memory & Cognition, 35*(3), 418–431. https://doi.org/10.3758/BF03193282

McRae, K., Cree, G.S., Seidenberg, M. S., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods, 37*(4), 547–559. https://doi.org/10.3758/BF03192726

Mcwilliams, J., & Schmitter-Edgecombe, M. (2008). Semantic memory organization during the early stage of recovery from traumatic brain injury. *Brain injury : BI, 22*(3), 243-253. https://doi.org/10.1080/02699050801935252

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. *Computer Science, 1-12.* arXiv preprint arXiv:1301.3781.

Pulvermuller, F. (2013). How neurons make meaning: brain mechanisms for embodied and abstract-symbolic semantics. *Trends in Cognitive Sciences, 17*(9), 458–470. https://doi.org/10.1016/j.tics.2013.06.004

Samson, D., & Pillon, S. (2003). A case of impaired knowledge for fruits and vegetables. *Cognitive Neuropsychology, 20*, 373–400.

Schacter, D. L., Dobbins, I. G., & Schnyer, D. M. (2004). Specificity of priming: a cognitive neuroscience perspective. *Nature Reviews. Neuroscience, 5*(11), 853–862. https://doi.org/10.1038/nrn1534

Scott, G. G., Keitel, A., Becirspahic, M., Yao, B., & Sereno, S. C. (2019). The Glasgow Norms: Ratings of 5,500 words on nine scales. *Behavior Research and Methods, 51*, 1258–1270. https://doi.org/10.3758/s13428-018-1099-3

Skelac, I., & Jandrić, A. (2020). Meaning as Use: From Wittgenstein to Google's Word2vec. In *Guide to Deep Learning Basics* (pp. 41–53). Springer. https://doi.org/10.1007/978-3-030-37591-1_5

Taylor, K. I., Devereux, B. J., & Tyler, L. K. (2011). Conceptual structure: Towards an integrated neuro-cognitive account. *Language & Cognitive Processes, 26*(9), 1368–1401. https://doi.org/10.1080/01690965.2011.568227

Toglia, M.P. (2009). Withstanding the test of time: the 1978 semantic word norms. *Behavior Research Methods, 41*(2), 531–533. https://doi.org/10.3758/BRM.41.2.531

Tyler, L. K., Chiu, S., Zhuang, J., Randall, B., Devereux, B. J., Wright, P., … Taylor, K. I. (2013). Objects and categories: feature statistics and object processing in the ventral stream. *Journal of Cognitive Neuroscience, 25*(10), 1723–1735. https://doi.org/10.1162/jocn_a_00419

Tyler, L. K., & Moss, H. E. (2001). Towards a distributed account of conceptual knowledge. *Trends in Cognitive Sciences, 5*(6), 244–252. https://doi.org/10.1016/s1364-6613(00)01651-x

Van Rensbergen, B., De Deyne, S., & Storms, G. (2016). Estimating affective word covariates using word association data. *Behavior Research Methods, 48*(4), 1644–1652. https://doi.org/10.3758/s13428-015-0680-2

Vigliocco, G., Vinson, D. P., Lewis, W., & Garrett, M. F. (2004). Representing the meanings of object and action words: The featural and unitary semantic space hypothesis. *Cognitive Psychology, 48*(4), 422–488. https://doi.org/10.1016/j.cogpsych.2003.09.001

Vinson, D. P., & Vigliocco, G. (2008). Semantic feature production norms for a large set of objects and events. *Behavior Research Methods, 40*(1), 183–190. https://doi.org/10.3758/brm.40.1.183

Wang, X., Men, W., Gao, J., Caramazza, A., & Bi, Y. (2020). Two Forms of Knowledge Representations in the Human Brain. *Neuron, 107,* 1-11. https://doi.org/10.1016/j.neuron.2020.04.010

Xiang, W., Lin, F., & Jiang, Z. (2015). The modeling and analysis of semantic features for the Chinese nouns. *Chinese Journal of Rehabilitation Medicine, 30*(11), 1118–1124.

Xie, W., Bainbridge, W. A., Inati, S. K., Baker, C. I., & Zaghloul, K. A. (2020). Memorability of words in arbitrary verbal associations modulates memory retrieval in the anterior temporal lobe. *Nature Human Behaviour, 4,* 937–948. https://doi.org/10.1038/s41562-020-0901-2

Yilmaz, S., & Toklu, S. (2020). A deep learning analysis on question classification task using Word2vec representations. *Neural Computing and Applications, 32*(7), 2909–2928. https://doi.org/10.1007/s00521-020-04725-w

Data availability statement

The datasets generated and analyzed during the current study are available in the supplementary materials (https://osf.io/ug5dt/). If you find any mistake in the dataset, we will appreciate it if you contact us to correct them.

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