

Database Design of an Online E-Learning Tool of Chinese Classifiers

Helena Hong Gao

Nanyang Technological University

14 Nanyang Drive, HSS-03-05

Singapore 637332

helenagao@ntu.edu.sg

Abstract

Chinese noun classifiers are an indispensable part of the Chinese language, but are difficult for non-native speakers to use correctly. Chinese language teachers often face challenges in finding an effective way to teach classifiers, as the rules for defining which nouns can be associated with which classifiers are not straightforward. Many theoretical studies have explored the nature of Chinese classifiers, but few studies take an empirical approach to the investigation of effective teaching and learning methods of classifiers. Learners often find that existing dictionaries either do not have classifiers as lexical entries, or give very brief explanations that are hardly helpful. This paper presents the progress of an ongoing project on the construction of an e-dictionary of Chinese classifiers. The objective of the project is to provide a platform for Chinese language learners to explore and learn classifier uses in a bottom-up fashion. The current work is on the design of an e-learning tool database and its connection to the e-dictionary database. Descriptions of the design and the functions of the e-learning tool are provided in the paper.

1 Introduction

As a classifier language, Chinese does not provide a way for its speakers to avoid using classifiers. That is, they are a compulsory grammatical element in a phrase structure. The basic construction of a classifier phrase has a numeral (e.g., *yī* ‘one’, *shí* ‘ten’), or a determiner (e.g., *zhè* ‘this’, *nà* ‘that’), or a qualifier (e.g., *jǐ* ‘several’, *měi* ‘each’), placed before a classifier and a noun after it. Thus, a simple English noun phrase such as ‘a book’ needs to be expressed in Chinese with the classifier *běn* in between the numeral *yī* and the noun, ‘book’ as *yībēn shū*. In brief, a classifier is a word or morpheme that is used to classify nouns based on their inherent semantic features. However, the semantics and the uses of Chinese classifiers have become far more complex than their syntactic structure looks. It is hard to define their lexical meanings and their uses seem to have rules to follow but violations are common. It is impossible for learners to make a correct choice of a classifier if they simply follow its grammatical rules. This is mainly because Chinese classifiers contain information about the features of the noun referents they can be associated with. However, most of the classifiers can be associated with a number of different types of nouns. Their noun-dependent meanings are inexplicit and ambiguous. Conventional dictionaries give brief definitions of classifiers, which is a way to avoid complicated descriptions and lengthy listing of their associated nouns, but are of little help to language learners.

Classifiers can be divided into different types based on their semantic functions. Some of them carry the unique features of the Chinese language; others are representational of classifier languages, and yet all of them have the functions of measure words. Regarding the differences between classifiers and measure words, Tai & Wang (1990) stated that “A classifier categorizes a class of nouns by picking out some salient perceptual properties, either physically or functionally based, which are permanently associated with entities named by the class of nouns; a measure word does not categorize but denotes the quantity of the entity named by a noun”. This definition makes a clear distinction between a classifier and a measure word from a multi-dimensional perspective. Measure words are language universal while classifiers are

This work is licensed under a Creative Commons Attribution 4.0 International Licence. Page numbers and proceedings footer are added by the organisers. Licence details: <http://creativecommons.org/licenses/by/4.0/>

language specific. There is an ontological base on which classifiers and nouns are associated with (Sowa, 2000; Huang and Ahrens, 2003; Philpot et al., 2003; Nichols et al., 2005), while measure word associations with nouns could be simply based on the notion of quantification. Understanding the differences between the two concepts can help learners of Chinese increase their awareness of the semantic and cognitive bases of classifier associations with nouns (Gao, 2010; Gao, 2011; Quek and Gao, 2011).

Due to the complexity of classifiers' functions, different definitions and classifications have been found. Some researchers define Chinese classifiers based on their grammatical functions. For example, Chao (1968) divided classifiers into nine categories. They are 'classifiers or individual measures', 'classifiers associated with v-o', 'group measures', 'partitive measures', 'container measures', 'temporary measures', 'standard measures', 'quasi-measures or autonomous measures', and 'measures for verbs of action'. His classification shows that he did not distinguish between classifiers and measure words. The advantages of such a classification are that it includes all the types of classifiers mentioned above and that classifiers' measuring function is emphasized. But a big disadvantage is that the embedded meanings of the specific noun classifiers and the ontological nature of the noun referents that classifiers are associated with are largely ignored. This way of classification may help beginning learners to understand the basic functions of Chinese classifiers, but will not help more advanced learners.

Yue (2009) took a different approach. He treated classifiers and measure words as quantifiers and divided those collected from corpus data into eleven categories based on the kinds of nouns the quantifiers are associated with. They were defined as quantifiers 'representing a group of people', 'indicating groups of animals', 'representing types', 'representing individual thing or person', 'representing a pair', 'representing a set', 'representing a huge amount of things', 'representing a slight amount of things', 'representing capacity', 'representing weather', and 'representing emotions'. Regardless of the unnecessary new term he used to refer to classifiers, his classification is more cognitively based and closer to language learners' knowledge of noun referents and their categories.

Using computer technology to apply empirical research findings of classifier knowledge to natural language processing (NLP) has provided a new approach for the semantic analysis of classifiers (Nirenburg and Raskin, 2004; Hwang et al., 2008) and for computer-assisted language learning (Guo and Zhong, 2005). However, no e-learning systems developed so far have been found to be able to help language learners to use the semantic features of classifiers' associated nouns to learn classifiers systematically. Yet, the emergence of Computer-Assisted Language Learning (CALL) has made it possible for language learners to explore various kinds of user-friendly and flexible e-learning tools (Davies, 2011). CALL incorporates technology into the language learning process and also applies itself across a broad spectrum of teaching styles, textbooks, and courses (Donaldson and Haggstrom, 2006). Its bidirectional and individualized features make it possible for learners to use it effectively to improve different aspects of language skills (Mallon, 2006; Chang et al., 2008).

The idea of designing an e-dictionary of Chinese classifiers is similar to that of CALL. Findings from empirical studies on classifier learning provide a practical guideline in the process of the designing. In order to make the e-dictionary a useful learning tool for both beginning and advanced learners of Chinese, measure words and classifiers are both labelled as classifiers. However, in the feature descriptions learners can understand and identify the functions of the words categorically.

Currently the dictionary database includes 859 classifiers collected from dictionaries and other resources. The number of associated nouns classified is currently 6420. Different tables (as sub-databases) are set up according to the classifications of classifiers and the nouns included. In addition to the conventional functions of a dictionary built up for the e-dictionary of classifiers, an e-learning system is implemented to allow learners at different levels to have a self-paced exploration of the relationships between a classifier and a noun or many nouns from different categories. In this paper the focus will be on the descriptions of the designs of the e-learning database and its interface.

2 Classifier-based Classification of Noun Categories

Classifiers must be used together with nouns to form classifier phrases but their associations with nouns are not contextually based nor are they of a free choice. The mapping can be complicated. A classifier

can be associated with a number of nouns from different categories and a noun can be mapped to more than one classifiers. For example, the classifier *tiáo* enters into nine noun categories and the noun *chē* 'car' can be associated with *liàng* and *tái* as well. Learners may quickly feel intimidated when at a first trial to identify the possibilities of the multi associations. Therefore, in designing the database, instead of mapping classifiers to nouns directly, we make use of classifiers' noun-dependent features to first identify all the nouns that each classifier can be associated with and then classify the nouns into categories. So far eleven noun categories have been identified and classified as 'nature', 'humans & body parts', 'animals', 'vegetables & fruits', 'man-made objects', 'buildings', 'clothing', 'food', 'furniture', 'tools', and 'vehicles'. A hierarchy of noun classifiers is built up according to the number of noun categories each classifier is associated with. These noun categories are not word classes defined with the principles in lexicology. They are defined based on the ontological categories of the noun referents of real-world entities, which are supposedly directly linked to learners' understanding of nouns and their referents of the language. Grouping classifiers' associated nouns into categories based on the ontological categories of noun referents is one of the special features of the design of this dictionary.

The classifiers are set in a hierarchical order in the database according to the number of noun categories they enter into. The highest number of noun categories that a classifier has been identified as being associable with more than one noun categories is nine. Of all the classifiers in the database, about more than 50% of them are associated with more than three noun categories. The fewer noun categories a classifier is associated with, the easier it is assumed to be for learners to grasp. For example, the classifier *liàng* occurs only in the category of vehicles, (e.g., car, lorry, bicycle, etc.). Learners generally do not confuse or misuse it for other types of nouns. Due to the differences in the multi-categorical associations, some classifiers are more commonly used than others. The nineteen classifiers listed in Table 1 are the ones that are associated with at least three noun categories and they are the most commonly used ones as well.

In the analysis of linguistic categories, a cognitive approach defines categories by groups of features and relationships within certain linguistic domains. The occurrence of a noun with a particular classifier in a phrase structure is dependent upon the categorical features of both the noun and the classifier. However, the embedded semantic networks of the categories are not obviously well connected, which is mainly due to the diachronic and sociolinguistic changes of the Chinese language. As a result, native speakers' categorization dependent on not only noun referents' intrinsic properties but also their functional and human perceptual ones. In other words, classifier and noun associations encode as well human cognitive understandings of the real world entities. The use of classifiers has thus been found changed over time. More noun and classifier associations are found to be possible cross-categorically. That is, one single classifier can associate itself with a number of nouns from different noun categories and similarly, one single noun can also be associated with not one but two or three classifiers. This cross-categorization extension complicates the classification of classifiers to a great extent.

Theoretically, it does not seem to be possible for linguists to build a meta-theory for a systematic organization of logically transparent classifier-noun categories and thus hard for lexicographers to find an effective way to illustrate the semantic relationships between classifiers and nouns. The main obstacle in classifier acquisition seem to be due to the fact that the the nature of the semantic meanings of classifiers is opaque. The complex classifier associations with nouns have consequently caused noun categorizations to be linguistically unconventional.

Classifier in Chinese	Classifier in Pinyin	Number of noun categories the classifier is associated with	Examples of nouns the classifier is associated with
条	<i>tiáo</i>	9 (nature, humans & body parts, animals, vegetables & fruits, buildings, clothing, food, vehicles, other man-made objects)	rainbow, leg, snake, cucumber, road, scarf, potato chip, boat, necklace
根	<i>gēn</i>	7 (nature, humans & body parts, vegetables & fruits, buildings, food, tools, other man-made objects)	stick, bone, banana, pillar, sausage, needle, ribbon
块	<i>kuài</i>	6 (nature, humans & body parts, clothing, food, tools, other man-made objects)	stone, scar, handkerchief, candy, eraser, soap
层	<i>céng</i>	5 (nature, humans & body parts, building, clothing, other man-made objects)	wave/fog, skin, building storey, curtain, paper
张	<i>zhāng</i>	5 (humans & body parts, food, furniture, tool, other man-made objects)	mouth, pancake, bed, bow, map
只	<i>zhī</i>	5 (humans & body parts, animal, clothing, vehicle, other man-made objects)	ear, tiger, sock, sailing boat, watch
粒	<i>lì</i>	4 (nature, vegetables & fruits, food, other man-made objects)	sand, cherry, rice, sleeping tablet
段	<i>duàn</i>	4 (nature, vegetables & fruits, building, other man-made objects)	wood, lotus root, city wall, iron wire
口	<i>kǒu</i>	4 (humans & body parts, animal, tools, other man-made objects)	person (people), pig, sword, well
面	<i>miàn</i>	4 (buildings, tools, furniture, other man-made objects)	wall, drum, mirror, flag
节	<i>jié</i>	4 (building, food, tool, vehicle)	chimney, sugar cane, battery, railway carriage
道	<i>dào</i>	3 (nature, humans & body parts, building)	lightening, eyebrow, dam
滴	<i>dī</i>	3 (nature, humans & body parts, other man-made objects)	water / rain, blood, ink
件	<i>jiàn</i>	3 (clothing, tools, other man-made objects)	shirt, (musical) instrument, toy
把	<i>bǎ</i>	3 (furniture, tools, other man-made objects)	chair, knife, cello
截	<i>jié</i>	3 (nature, tools, other man-made objects)	rope, pencil, pipe
颗	<i>kē</i>	3 (nature, humans & body parts, other man-made objects)	star, tooth, artillery shell
片	<i>piàn</i>	3 (nature, food, other man-made objects)	leaf, loaf, tablet
枝	<i>zhī</i>	3 (nature, tools, other man-made objects)	rose, pen, arrow / rifle

Table 1: Contents of the main database for the e-dictionary. Each role in the table is a sub-database in the system.

Studies show that native speakers of Chinese tend to take a cognitively-based bottom-up approach as a strategy to the learning of classifiers while second language learners of Chinese tend to take a top-down approach but often find their learning outcome inefficient (Soh and Gao, 2009; Gao, 2010; Quek and Gao, 2011). The cognitive approach taken for the design of the database is based on the findings of empirical studies on Chinese classifier learning by adults and children of both native and non-native speakers of Chinese. The classifier-based classifications of noun categories that reflect the ontological knowledge of this category of linguistic terms and its structure are assumed to be able to activate learners' cognitive processes when exploring the pragmatic use of classifiers.

3 Noun-based Semantic Features of Classifiers Decomposed

Table 1 is an illustration of the contents of the main database for the e-dictionary. Each role in the table is a sub-database in the system.

Table 1 is a demonstration of the semantic features of some most commonly used noun classifiers and their associated nouns. Through semantic decomposition of the noun-based classifier features, the cognitive mapping between a classifier and its associated nouns are revealed. Take the classifier *tiáo* for example (see Figure 1). It is associated with nouns such as rainbow, leg, snake, cucumber, road, scarf, potato chip, boat and necklace, which are from nine of the eleven noun categories listed in Section 2. Despite of the different categories they belong to, the nine nouns share one same property — the shape of the noun referents that is defined as 'longitudinal'. This shows that the classifier *tiáo* is inhabited with this semantic feature and it is possibly the cognitive basis on which native speakers of Chinese associate it with the related noun referents accordingly.

Similarly, the classifier *gēn* is used with the nouns such as stick, bone, banana, pillar, sausage, needle, and ribbon that belong to seven noun categories respectively. These nouns possess the same 'longitudinal' feature as *tiáo* does. This means that extracting one same feature from *gēn* and *tiáo* is not helpful enough for learners to understand the differences between the two classifiers though classifying nouns into ontological categories can constrain the interference to learners to a certain extent. What needs to be further specified is to define each noun with a unique feature of its own, no matter whether it is from its lexical semantic meanings, pragmatic functions, or human perceptions. For example, in addition to the feature labelled as 'longitudinal', 'for supporting walking' is added as a feature to 'stick', 'a piece of human skeleton' to 'bone', 'turning from green to yellow when ripe' to 'banana', 'one end stuck to the ground' to 'pillar'. More specifications are needed until finally each noun is distinguished from other nouns that are associated with one same classifier. These definitions are the core part of the database in the e-learning tool system linked to the e-dictionary.

4 Methodology

4.1 Application of Cognitive Strategies in Noun Classifier Acquisition

In this section we describe an approach that is used for extending the design of the e-dictionary to that of an e-learning tool as another part of the project. Developed first in the software environment of FileMaker Pro 8.5 (see Figure 2), the dictionary is established on a database system. Categorical records created as data files are used to store the associated nouns. The records created so far include eleven categories of nouns as are described in Section 2. Such a categorization appears explicit, but its top-down approach fails to reveal the feature-based mapping between a classifier and its associated nouns. The objective of the e-learning approach, on the other hand, is to guide users to search for correct classifier and noun pairs by looking for the defined features of the noun referents, firstly from those broadly defined as 'animacy', 'shape', 'size', 'thickness', 'length', and 'function' to those specific ones extracted from each particular noun referent.

With such a bottom-up approach, the e-dictionary allows users to learn to use the particularly inter-related features of a classifier and its associated noun referents in a case-by-case fashion. In this way learners can better understand the point that a classifier reflects the cognitive classification of its associated noun referents. Each individual record thus contains both general and specific information of a classifier and its associated nouns as data entries, The features decomposed from the noun referents are

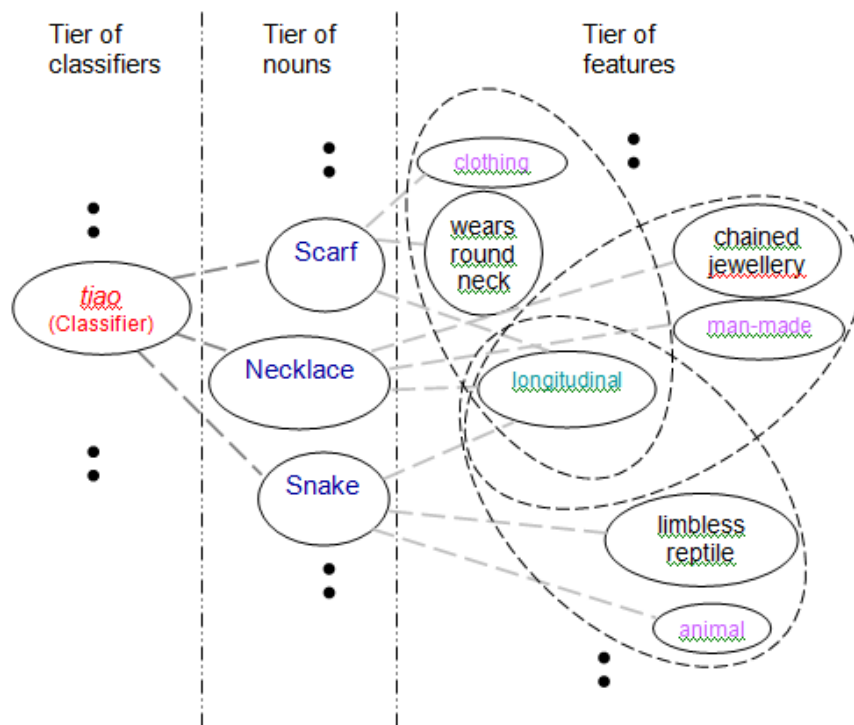


Figure 1: Mapping among the tiers of classifiers, nouns, and defined features.

defined and recorded as independent data entries linked to the e-learning tool. For example, if a learner wants to know which classifier is the correct one for ‘boat’, he or she can enter the word ‘boat’, find its category as ‘vehicles’, and choose its shape as ‘longitudinal’. Then *tiao* should automatically pop up in this case because ‘boat’ is the only noun referent from the ‘vehicles’ category (see Table 2). In other cases where there are two or more noun referents that are featured as ‘longitudinal’, the user will be guided to look for a more specific or unique feature with a few more clicks on the users’ interface.

The e-learning environment in the dictionary also provides users the classifier phrases that are commonly used but they may not be easy for learners to acquire. Take the classifier *zhī* for example. It is associated with noun referents that belong to ‘animals and body-parts’, and ‘man-made objects’, such as ‘bird’, ‘hand’, and ‘pen’. The unique perceptual features of these noun referents are identified and built into the e-learning system so that users can click different categories on the interface to make particular associations as long as they have some general knowledge of the noun referents in terms of functions and perceptual features.

CL in Character	CL in Pinyin	Associated nouns in Chinese	Associated nouns in English	Associated noun categories	Shape
根	<i>gēn</i>	电线杆	telegraph pole	buildings	longitudinal
根	<i>gēn</i>	骨头	bone	humans & body parts	longitudinal
根	<i>gēn</i>	棍	stick	tools	longitudinal
根	<i>gēn</i>	黄瓜	cucumber	vegetables & fruits	longitudinal
根	<i>gēn</i>	面条	noodle	food	longitudinal
根	<i>gēn</i>	绳子	rope	tools	longitudinal
根	<i>gēn</i>	丝带	ribbon	other man-made objects	longitudinal

Table 2: An example of how nouns are grouped in the database.

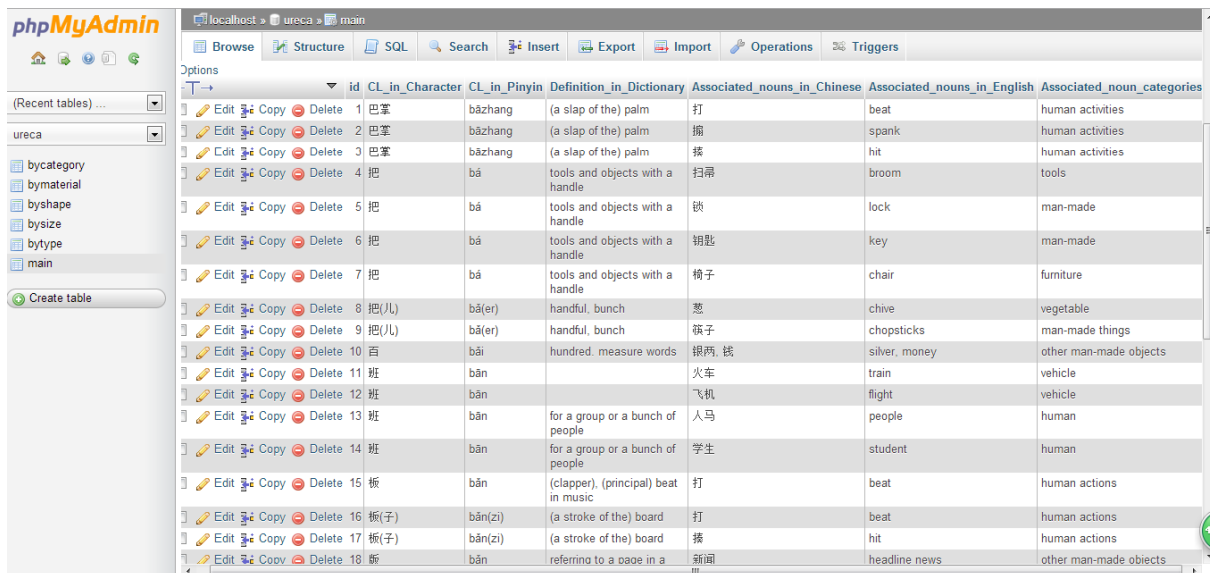


Figure 2: A view of the database interface.

4.2 Database Construction Using MySQL

In a Windows web development environment WAMPSEVER, this database is created under a database management system MySQL. Compared with other database systems, MySQL is relatively more reliable and easier to use, especially for the design of web applications. MySQL database can also be handled and managed using tools like phpMyAdmin. Figure 2 is a display of the database in the web environment.

Data in MySQL are stored in different tables and every unit in each table can be referred to by its row and column index. This feature makes online search convenient and applicable to the design of web application.

As shown in Figure 2, there are six tables that have been created in the database. In table ‘main’, data are sorted by classifiers. Basically, this table contains all of the information in a conventional dictionary that takes classifiers as lexical entries. This table is mainly used for searching classifiers for a noun or searching nouns for a classifier. Users can conduct their search for nouns using both Chinese and English. Search in Chinese has been designed to allow input either in Chinese characters or in Pinyin for both classifiers and nouns. The definition of a classifier, its associated nouns in Chinese and English, and the categories that its associated nouns belong to can all be searched categorically. The search outcome is then presented in a result page.

In table ‘byshape’, data are sorted by the shapes of noun referents such as ‘longitudinal’, ‘rectangular’, and ‘round’. In the other tables specific features of noun referents that have been so far defined are sorted respectively by size, quality of material, and ontological categories. In these tables, not every noun in table ‘main’ is included, as some nouns cannot be described using these features. For example, the word ‘customer’ cannot be described under the feature of shape, but it is classified into the category of human. Hence, it exists in table ‘bycategory’ but not in the ‘byshape’ one.

The database is designed in such a way so as to increase the efficiency of the search function in the web application. The idea of implementing this searching function is to add the conditions one by one so as to narrow down the search field. For example, ‘hair’ is in the category of ‘human & body parts’ with features ‘longitudinal’, ‘thin’ and ‘soft’. Users can conduct the following step-by-step search:

1. Search in table ‘bycategory’ to find all the nouns in the category of ‘human & body parts’. Call them Group 1.
2. Search nouns with shape ‘longitudinal’ and in Group 1 in table ‘byshape’. Call them Group 2.
3. Search in table ‘byshape’ for the nouns with the condition ‘thin’ and in Group 2. Call them Group 3.



(a) Input page.



(b) Drop-down list search page.

Figure 3: The web application user interface.

4. Search in table 'bymaterial', for the nouns with the condition 'soft'.

At Step 4 when the condition 'soft' is chosen, users should be left with the only noun 'hair'. As the sizes of Groups 1, 2 and 3 become smaller, the search time is reduced. If more features are added, then the steps of search can be repeated until the target noun is found.

4.3 Design of the Web Application Interface

The web application interface is designed using PHP, a server-side scripting language. Basically, it is a dynamic web page connected to the MySQL database built up for this purpose. It means that the content of this web application depends on the database and what is submitted to the server. On the client side or the web application interface for users, as shown in Figure 3a and Figure 3b, HTML language is applied to build up the basic structure as well as the presentation of the website.

The method used in the design of the web application interface make the e-learning tool instructive and self-exploratory. Once a user clicks one of the three links on the side bar, he or she will be directed to a webpage shown in Figure 3a or Figure 3b, where the search function is contained in an HTML

form. With HTML form attribute 'action', data will be submitted to a specific page, that is, the php script containing that search function.

Inside the respective php script, there will be several common command lines:

```
@ $db = new mysqli('localhost', 'root', '', 'ureca');
$db->set_charset("utf8");
if (mysqli_connect_errno()) {
    echo 'Error: Could not connect to database.';
    echo 'Please try again later.';
    exit; }
}
```

These commands are to connect to and set the character set of the database built up for this purpose. Here *utf8* is used as there are both English and Chinese characters in the database. On these three web pages, forms are submitted using the 'post' method. POST method sends form-data as HTTP (Hypertext Transfer Protocol) post transaction and has no limitations on the size of data. However, restrictions may exist due to the nature of the database. Search results will then be grouped and displayed, with searchable words hyperlinked. Here the method used is GET, which works better for non-secure data with a limited size.

This web application interface enables learners to discover a noun classifier in three types of search, 'searching a noun for its classifier(s)', 'searching a classifier for its associated noun(s)', and 'searching classifiers for a group of nouns by restricting one or more conditions'.

- Search type 1: *Searching a noun for its classifier(s)*.

This function is similar to any other online dictionaries. Learners insert a noun to do a simple and direct search. On the result page learners can see which classifier or classifiers can be applied to the noun and the definitions of the classifier(s). This is an early design for the e-dictionary. An example of the steps is given in Figure 3b.

- Search type 2: *searching a classifier for its associated noun(s)*.

This function is for learners to start their search with a classifier. It is assumed that learners had learnt a classifier but had not known yet what nouns could be used with the classifier. The search result is shown on a new page that includes the definition of the classifier and its associated nouns.

The difference between search types 1 and 2 is that with search type 1 the result is simple and direct. Learners can understand right away how to form a classifier phrase with the result given. The result of search type 2, however, can display all the nouns that a single classifier can be associated with. Learners of Chinese at the beginning stage may feel intimidated seeing the result showing more nouns than they expect as they may not have learned yet why these different nouns are all related to each other.

- Search type 3: *searching classifiers for a group of nouns by restricting one or more conditions*.

Instead of keeping all the conditions as the default in search types 1 and 2, learners can choose one condition or more at a time from the drop-down lists. The chosen lists allow the system to fetch desired data directly from the databases. Learners can delete any of the chosen conditions to start a new search and to compare the results. This function is for more advanced learners who have learned the general principles of classifier-noun associations and who have a clear target in their search for a particular type of classifiers or nouns.

Further search is also designed in the primary result page. This function is currently shown as a hyper-link, which is dynamically generated and assigned with a value in advance. The varieties of search functions and illustrations are expected to eventually enhance learners' understanding of the multi-dimensional noun-classifier associations.

5 Discussion

What is presented in this paper is the progress of an on-going project on the building up of an e-learning tool for learning Chinese classifiers. The aim of this project is to clarify the embedded relationships

between classifiers and their associated nouns so as to assist Chinese language learners in the acquisition of classifier phrases. At the first stage of the project, classifiers and their associated nouns were collected from dictionaries and other resources. A database for the e-dictionary part which contains the functions that are characteristic of conventional dictionaries was designed and set up. Learners can search for a classifier or a noun for their association as a classifier phrase. Then, a feature-based approach in designing the classifier e-dictionary was extended to an e-learning environment created for learners to explore. At the current stage, the task is on the design and setting up of an e-learning system attached to the e-dictionary. As an experiment, MySQL was used to build up the database shared by the conventional dictionary and the e-learning system.

The structure of this database is formed in the way in which classifiers and nouns are stored in different tables but can be linked together. All the information stored in different tables was connected through respective grouping criteria which allow the data to be extended to the e-learning environment. One table contains one type of information, such as classifiers in character, classifiers in pinyin, definitions of classifiers, types of classifiers, classifier associated nouns in Chinese and English, categories of the associated nouns, and semantic features of the nouns. Every two tables share at least one common parameter, which enables cross-table search as described in Section 4.2. Such a design is able to boost the efficiency of the search function. In addition, the database can be enriched easily through MySQL code or phpMyAdmin to import new data.

A web application for self-learning in the e-learning environment was designed using PHP language. It serves as an e-learning tool for learning Chinese classifiers. The various searching functions provide progressive search for specified features of classifiers' associated nouns and their classified categories.

The feasibility of the functions of the e-learning tool and its web application need to be further improved. Currently there are a few limitations. For example, subjectivity is a limitation of this database. In the process of decomposing nouns into respective semantic features, human cognition plays an important role. However, this parameter varies from person to person. Moreover, speakers of Chinese in different regions may tend to use different cognitive strategies in their associations with the semantic meanings of classifiers. Therefore, data from experiments and empirical studies are needed for the future improvement of the semantic analysis and descriptions of the noun-classifier associations. Another aspect to improve is that regional featured uses of certain classifiers such as Singaporean Chinese speakers' use of *li* with noun referents that are both big and small (e.g., 'watermelon' and 'bean') can be explained and included in separate tables so that learners can be aware of the regional differences in classifier use.

The advantages of the web application design are its multipurpose search functions and flexible links to the various parts of the database behind. With the various search functions, learners will be able to investigate classifiers from different aspects, which is ideal for self-learning. On the page of 'Searching a noun for its classifier(s)', both Chinese and English entries are acceptable, which makes it easier for learners to explore and make a flexible use of its learning functions. However, on the page of 'Searching a classifier for its associated noun(s)', only Chinese character entries are available. A future addition can be made to allow entries by Pinyin as well.

For the database development in the future, the web application is designed to be linked to the database in a dynamical fashion. Any changes made to the database can be reflected on the web page automatically. This will also allow us to make further development without much of a change in the current layout in the e-learning system.

6 Conclusion

Based on the Chinese classifier e-dictionary of (Gao, 2011), designed to help students learn the proper use of Chinese classifiers, this paper further explores the designs of the database and of an e-learning tool interface to better understand the association of classifiers and nouns. In this experimental version of the e-learning tool design, 859 classifiers and 6420 associated nouns were stored and classified in different tables according to the respective noun referents' semantic features and prominent cognitive features. The system built-up with MySQL has shown its convenient linkage to database management tool phpMyAdmin and web-design language PHP. As the base of the e-learning tool, the database with an

interface built-in can be searched step by step with individual or combined functions. The results can be displayed on the users' webpage. Learners can examine the property of a classifier and the link between this classifier and its associated nouns from several perspectives via various search functions. The multi-functional feature of this webpage is the design of the drop-down list search, which allows users to discover classifiers' noun-dependent features case by case. To make use of the advantage of internet, further investigation of another feature within a search can be made possible through hyperlinked text.

The final goal is to make the outcome of this project available online as learning resources for the general public and as an e-learning tool for Chinese language learners. Further development of this project and explorations of other possible database designs are necessary as our end goal is to provide an effective learning tool. Experimental studies are also needed to discriminate the subjectivity of the descriptions of human cognition in the illustrations of classifier-noun associations.

Acknowledgements

I would like to thank my research students Ong Shi Lun for the data entry, and Chen Jiaying for her work on the database trials.

References

- Yu-Chia Chang, Jason S. Chang, Hao-Jan Chen, and Hsien-Chin Liou. 2008. An automatic collocation writing assistant for Taiwanese EFL learners: A case of corpus-based NLP technology. *Computer Assisted Language Learning*, 21(3):283–299.
- Yuen Ren Chao. 1968. *A Grammar of Spoken Chinese*. University of California Press.
- Graham Davies. 2011. Introduction to multimedia CALL. In Graham Davies, editor, *Information and Communications Technology for Language Teachers (ICT4LT)*. Thames Valley University, Slough. www.ict4lt.org/en/en_mod2-2.htm.
- Randall P. Donaldson and Margaret A. Hagstrom. 2006. *Changing Language Education Through CALL*. Routledge.
- Helena Hong Gao. 2010. A study of the Swedish speakers' learning of Chinese classifiers. *Nordic Journal of Linguistics*, 33:56–63.
- Helena Hong Gao. 2011. E-learning design for Chinese classifiers: Reclassification of nouns for a novel approach. In R. Kwan, C. McNaught, P. Tsang, F. Lee Wang, and K. C. Li, editors, *Communications in Computer and Information Science (CCIS)*, number 177, pages 186–199. Springer-Verlag.
- Hui Guo and Huayan Zhong. 2005. Chinese classifier assignment using SVMs. In *4th SIGHAN Workshop on Chinese Language Processing*, pages 25–31, Jeju, South Korea.
- Chu-Ren Huang and Katherine Ahrens. 2003. Individuals, kinds and events: Classifier coercion of nouns. *Language Sciences*, 25:353–373.
- Soonhee Hwang, Ae-Sun Yoon, and Hyuk-Chul Kwon. 2008. Semantic representation of Korean numeral classifier and its ontology building for HLT applications. *Language Resources and Evaluation*, 42:151–172.
- Adrian Mallon. 2006. ELingua Latina: Designing a classical-language e-learning resource. *Computer Assisted Language Learning*, 19(4):373–387.
- Eric Nichols, Francis Bond, and Daniel Flickinger. 2005. Robust ontology acquisition from machine-readable dictionaries. In *19th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 1111–1116, Edinburgh, UK.
- Sergei Nirenburg and Victor Raskin. 2004. *Ontological Semantics*. MIT Press.
- Andrew G. Philpot, Michael Fleischman, and Eduard H. Hovy. 2003. Semi-automatic construction of a general purpose ontology. In *Proceedings of the International Lisp Conference*, pages 1–8, New York, NY, USA.
- See Ling Quek and Helena Hong Gao. 2011. An experimental investigation of the cognitive basis of Malaysian Chinese speakers' association of one noun with multiple classifiers. In *12th Chinese Lexical Semantics Workshop*, pages 232–243, Taipei, Taiwan.

- Ning En Christabelle Soh and Helena Hong Gao. 2009. Chinese noun classifier usage by Singaporean bilingual children. *Proceedings of the URECA@NTU*, pages 88–91.
- John F. Sowa. 2000. *Knowledge Representation*. Brooks Cole Publishing Co.
- James H.-Y. Tai and Lianqing Wang. 1990. A semantic study of the classifier Tiao. *Journal of the Chinese Language Teachers Association*, 25(1):35–56.
- Weiwei Yue. 2009. Contrastive analysis of quantifiers in Chinese and English from a cognitive perspective. Master's thesis, Shandong Normal University.