Saizhu Hu

60461

Lost and added in translation:
A corpus-based study on Chinese numeral classifiers in translation between Chinese and English

Utracone i dodane w tłumaczeniu:
Analiza korpusowa chińskich klasyfikatorów liczbowych w tłumaczeniu pomiędzy jęz. chińskim a jęz. angielskim

Poznań, 2023

ACKNOWLEDGEMENTS

This project would never have been completed without the help and support of many people.

First, I would like to express my deepest gratitude to my supervisor Prof. Marcin Kilarski. This endeavour would not have been possible without your valuable and insightful comments and your step-by-step and patient guidance. Your experience and practical advice have helped me tremendously at all times in my research and writing of this dissertation.

I would also like to extend my gratitude to Prof. Haitao Liu of Zhejiang University. Thank you for all the suggestions as to the choice of the topic, the research methods concerning data processing and analysis, and the comments on my papers and research in general.

Special thanks should also go to Prof. Joanna Pawelczyk and Prof. Bogusława Whyatt. You have given me a lot of help in coordinating my doctorate studies as I took an individual program and had to travel between Poland and China during my doctorate studies. Furthermore, you have given me so much warmth and encouragement during my stay in Poland.

Next, I am grateful to all the teachers, colleagues and students in Faculty of English, Adam Mickiewicz University. I had the pleasure of attending your classes and activities, and your friendliness and helpfulness are always impressive to me.

I'd also like to acknowledge the support from NingboTech University and my colleagues and students there. Thank you for allowing me to leave for my doctorate studies and for sharing my duties and responsibilities in the process so that I can be more devoted to my project.

Lastly, I have to thank my friends and family. Your unconditional love and support have always been the source of my power to move on in this difficult process. Sorry that I had to excuse myself from so many gatherings and put you to take so much of my anxiety and pains in these years.

OŚWIADCZENIE

Ja, niżej podpisany/a

Saizhu Hu

···

przedkładam rozprawę doktorską

pt. Lost and added in translation: A corpus-based study on Chinese numeral classifiers
in translation between Chinese and English

Utracone i dodane w tłumaczeniu: Analiza korpusowa chińskich klasyfikatorów
liczbowych w tłumaczeniu pomiędzy jęz. chińskim a jęz. angielskim

···

na Uniwersytecie im. Adama Mickiewicza w Poznaniu
i oświadczam,
że napisałem/am ją samodzielnie.

Oznacza to, że przy pisaniu pracy, poza niezbędnymi konsultacjami, nie korzysta-
łem/am z pomocy innych osób, a w szczególności nie zlecałem/am opracowania
rozprawy lub jej istotnych części innym osobom, ani nie odpisywałem/am tej roz-
prawy lub jej istotnych części od innych osób.

Jednocześnie przyjmuję do wiadomości, że gdyby powyższe oświadczenie oka-
zało się nieprawdziwe, decyzja o wydaniu mi dyplomu zostanie cofnięta.

May 10, 2023	Saizhu Hu
(miejscowość, data)	(czytelny podpis)

Table of contents

List of tables

List of figures

List of Abbreviations

A	subject of a transitive verb
ACC	accusative
ADJ	adjective
AGT	agent
ANIM	animate
ART	article
AUG	augmentative
AUX	auxiliary
CJ	conjoint
CLF	classifier
CN	Chinese
CONJ	conjunction
DEM	demonstrative
DET	determiner
DIR	directional
EN	English
FEM F.	feminine
FOC	focus
HUM	human
LOC	locative
MASC M.	masculine
MENS	measure word
MOD	modifier
N	noun
NCL	noun class
neut.	neutral
NUM	numeral
OBV	obviative
PAST	past
PL	plural

PRON   pronoun

POSS   possessive

PRT   particle

SG   singular

Chapter 1: Introduction

1.1. Background

Languages are diverse in lexical choices and grammatical structures, and thus they may express equivalent meanings with different forms. Analytic languages have words with few morphemes and tend to be economic in morphological markers. Therefore, they rely more on free forms to express both lexical and grammatical meanings. For example, Mandarin Chinese uses independent words to express gender and definiteness. In contrast, synthetic languages have words containing more morphemes, so they use more morphological markers to express grammatical meanings, such as gender, number and case. Such grammatical categories provide another example of difference among such languages as French and Mandarin Chinese. For example, gender, number, and case occur in French but not in Mandarin Chinese, while numeral classifiers occur in Mandarin Chinese but not in French. Similarly, numeral classifiers do not occur in English, in which different forms express the meanings conveyed by numeral classifiers in Mandarin Chinese.

Numeral classifier systems are mostly found in East and Southeast Asia, with more isolated cases found in the languages of the Americas and Africa (Gil 2013). In China, numeral classifiers occur in all the main varieties of Chinese, including Mandarin, Wu, Min and Cantonese. This study focuses on numeral classifiers in Mandarin Chinese as the variety with the longest tradition of descriptive study.[1] Chinese has a fairly large number of at least a hundred numeral classifiers (e.g., Huang and Ahrens 2003; Gao and Malt 2009; Ma 2015).

---

[1] Unless specified explicitly, the term "Chinese" will therefore be used with reference to Mandarin Chinese.

Numeral classifiers can be independent words adjacent to numerals or quantifiers, and affixes or clitics attached to or fused with numerals (Aikhenvald 2017). In Chinese, numeral classifiers are independent elements occurring "with a number and/or a demonstrative, or certain quantifiers before a noun" (Li and Thompson 1981: 104). Numeral classifiers are obligatory elements in numeral noun phrases which enable the enumeration and categorization of noun referents in terms of, e.g., animacy, humanness and shape. Two types are commonly distinguished: sortal classifiers[2] indicating inherent properties of noun referents, and mensural classifiers denoting contingent quantities of noun referents (Her 2012: 402). Numeral classifiers can be used to individuate nouns and denote semantic properties of noun referents, and they can also be used to anaphorically track noun referents in discourse (Contini-Morava and Kilarski 2013). There is usually one general classifier and more specific classifiers in a classifier language. For example, in Chinese, there is one general classifier *gè* 'CLF:GENERAL' and a variety of specific classifiers, such as *tiáo* 'CLF:SLENDER' and *wèi* 'CLF:INDIVIDUAL, RESPECT'.

Chinese numeral classifiers have received substantial attention from linguists around the world especially in recent decades. Their obligatory occurrence in numeral noun phrases has long been the focus of interest typically in Chinese linguistics. Some linguists equal numeral classifiers with grammatical forms such as number markers (e.g., Her and Chen 2013) and determiners (e.g., Cheng and Sybesma 2012a). Since the 1970s, numeral classifiers have been increasingly recognized as semantically and pragmatically motivated and as important devices to categorize noun referents (e.g., Aikhenvald 2000). The interest in the semantic functions of numeral classifiers has been extended to the other elements in classifier structures and the effect of the presence and ellipsis of numeral classifiers on the interpretations of these elements and the classifier phrases as a whole (cf. Cheng and Sybesma 2012b; Cheng et al. 2017).

However, relatively little research has been done on the comparison of numeral classifiers with their corresponding forms in English translation, particularly with regard to the functionality of numeral classifiers. Such comparisons can directly show how different forms in a classifier and non-classifier language are used to express the same meanings, and how the presence and absence of numeral classifiers affect the interpretations

---

[2] In this dissertation, the term 'numeral classifiers' refers to sortal classifiers, while the term 'measure words' refers to mensural classifiers in Chinese.

of related elements and structures. Therefore, this study is meant to fill in this gap of research.

1.2. Aims

This study aims at a corpus-based analysis of Chinese numeral classifiers in their English translation. Chinese numeral classifiers were compared with corresponding forms in English in parallel corpora to show how a classifier language and a non-classifier language express lexical meanings, such as the semantic features of nouns referents, and grammatical meanings such as number, definiteness, and specificity. Contrastive studies of the use of numeral classifiers and corresponding forms were thus conducted based on parallel corpora to examine the semantic contribution of numeral classifiers to noun phrases, the semantic functions of numeral classifiers and their representation in English translation, and the discourse functions of numeral classifiers and their representation in English translation. More specifically, the study addresses the following research questions:

1) To what extent are Chinese numeral classifiers used with and without adjectives in Chinese?
2) How do Chinese numeral classifiers collocate with different types of nouns?
3) How do Chinese numeral classifiers and other elements contribute to the semantics of noun phrases?
4) What forms and elements in English correspond to Chinese numeral classifiers based on their lexical and grammatical features and semantic functions?
5) How are the discourse functions of Chinese numeral classifiers represented in English translation?

Based on the above research questions, the hypotheses are formulated as follows. Concerning the semantic contribution of Chinese numeral classifiers, general classifiers are used more frequently without adjectives. However, specific classifiers are more likely to occur with adjectives. Second, Chinese numeral classifiers are more likely to collocate with countable nouns instead of uncountable nouns. Third, numeral classifiers contribute to noun phrase reference with other elements in numeral noun phrases. As regards the translation of Chinese numeral classifiers based on their lexical and grammatical features, while Chinese numeral classifiers tend to be omitted in the translation into English,

definiteness expressed by Chinese numeral classifiers tends to be conveyed by articles in English. Therefore, the numeral *yī* 'one' is more likely to be translated into articles expressing definiteness rather than the numeral *one* only expressing specificity. Semantic properties denoted by Chinese numeral classifiers tend to be expressed by nouns in English. With regard to the English representation of the semantic functions of Chinese numeral classifiers, semantic units created by numeral classifiers tend to be reflected in singular and plural forms of countable nouns in English, but in measure words in English when their head nouns are uncountable. In contrast, semantic properties attributed to noun referents by Chinese numeral classifiers tend to be omitted in English translation. Concerning the translation of Chinese numeral classifiers with regard to their discourse functions, specific classifiers tend to be used to identify and recategorize referents, while general classifiers are more likely to manage reference in discourse. In managing referents in discourse, definiteness expressed by Chinese numeral classifiers tends to be expressed by articles in English. However, properties identified and recategorized by numeral classifiers are likely to be expressed by nouns in English.

These findings can contribute to studies in various fields. First, they can contribute to the studies on the typology of nominal classification, the distance of numeral classifiers from canonical and noncanonical gender (Corbett and Fedden 2016; Fedden et al. 2018), and language complexity with regard to multi-functionality of numeral classifiers, typically the notion of hidden complexity proposed by Bisang (2014). Secondly, comparing Chinese numeral classifiers with their corresponding forms in non-classifier languages based on their functionality can contribute to the studies on the acquisition and translation of numeral classifiers in applied linguistics. Third, the study can also help explore the cognitive mechanisms of language processing in classifier and non-classifier languages in neurolinguistics, as cognitive processing of numerals and nouns can be very different in the contexts with and without numeral classifiers.

1.3. Methodology

In order to address the above research questions, both quantitative and qualitative studies were conducted. Quantitative studies were carried out based on two self-compiled specialized parallel corpora of Chinese numeral classifiers based on the corpora of BCC (Xun

et al. 2016). The two corpora include 6700 pairs of Chinese-English numeral noun phrases without adjectives and 523 pairs of Chinese-English numeral noun phrases with adjectives. Data in the two corpora do not show in which direction numeral noun phrases are translated, as the online version of BCC, on which Corpus 1 and Corpus 2 were based, does not provide the source of its data or indicate the source languages of translation. Quantitative studies focused on the frequencies of different types of Chinese numeral classifiers collocated with other elements, e.g., adjectives and nouns, the comparison of different types of Chinese numeral classifiers with their equivalents in English in terms of their lexical and grammatical meanings and their semantic functions, and the issues concerning their semantic restrictions and contributions in noun phrases. Qualitative studies were made based on a parallel corpus of about 645 pairs of noun phrases in more than 411 pairs of sentences derived from five chapters of *The Three-Body Problem*, a Hugo Award science fiction novel by Liu Cixin (Liu [2008] 2014). Qualitative studies devoted more attention to issues concerning the discourse functions of Chinese numeral classifiers and their representation in English by examining how noun referents are identified, maintained and recategorized in discourse. A more detailed description of data collection and analysis is given in Chapter 4.

1.4. Structure of the dissertation

The remaining chapters of the dissertation are organized as follows. Chapter 2 gives a general sketch of the two main types of nominal classification systems, i.e., classifiers and gender. The distinction between numeral classifiers and other nominal classification devices, including other types of classifiers, is essential in this study, so particular attention is devoted to the typologies of nominal classification systems based on their morpho-syntactic expression and the degree of grammaticalization. Chapter 3 focuses on Chinese numeral classifiers in terms of their semantic and syntactic features and functions to serve as a foundation for the following corpus-based studies. Different approaches to Chinese numeral classifiers in Western and Chinese linguistics are also described in this part. The methodology of the study is discussed in Chapter 4. Chapter 5 examines the semantic contribution of different types of Chinese numeral classifiers and other elements of the classifier phrases, and Chapter 6 and Chapter 7 offer corpus-based Chinese-English

comparisons of numeral classifiers based on their functionality. Chapter 6 focuses on the comparison of Chinese numeral classifiers and other elements in noun phrases with their corresponding forms in English with regard to their semantic functions based on Corpus 1 and Corpus 2. Chapter 7 devotes more attention to the English translation of Chinese numeral classifiers with regard to their discourse functions based on Corpus 3. The study closes with a summary of the main findings in Chapter 8 including also a discussion of suggestions for further studies.

Chapter 2: Nominal classification systems

2.1. Introduction

Nominal classification, a classification of noun referents, is a pervasive feature in the languages of the world. Systems of nominal classification can be used to express a variety of features of noun referents, e.g., animacy, sex, humanness, and physical properties. Nominal classification has two major types: gender and classifiers. Gender, also referred to as noun classes, occurs in many Indo-European languages and languages of Africa, while classifiers, typically numeral classifiers, are widespread in the languages of East and Southeast Asia and Oceania.

Before Chinese numeral classifiers are examined in Chapter 3, this chapter is devoted to a description of nominal classification in general. In §2.2, I will discuss criteria that have been proposed to distinguish between gender and classifiers, and other types of categorization. In §2.3 and §2.4, I will focus on gender and classifiers, respectively, including such issues as their expression, semantic organization and functions. A brief account of the diachrony of nominal classification will be provided in §2.5 and the studies on nominal classification systems will be reviewed in §2.6. Finally, conclusions will be given in §2.7.

2.2. Nominal classification and other types of categorization

There is considerable variation in terms of the classification of nouns and their referents. The term 'nominal classification' is more frequently used in the literature (e.g., Mithun

1986; Grinevald 2000; Senft 2000; Aikhenvald 2004; Seifart 2010; Kilarski 2013; Passer 2016), while a number of other terms are also used interchangeably with 'nominal classification', such as 'noun categorization' (e.g., Craig 1986b; Aikhenvald 2000, 2017) and 'noun classification' (e.g., Lyons 1968; Allan 1977; Becker 1986).

Following Contini-Morava and Kilarski (2013: 265), the term 'nominal classification' is here used as a cover term to refer to "classification of nouns and/or extralinguistic entities to which nouns refer that is grammaticalized to some degree, and expressed in one or more syntactic contexts that relate to nouns", while the term 'noun classification' is used with reference to "systems that include at least some classification of nouns as linguistic forms". In this way, the use of the term 'nominal classification' reflects common semantic functions of all forms of nominal classification, and at the same time, shows commonalities in their expression (Lucy 2000: 331), while the term 'noun classification' is reserved for those classification systems based on features of nouns (Lucy 2000: 331; cf. Contini-Morava and Kilarski 2013: 265).

The distinction between nominal classification and noun classification reflects an ambiguity with regard to the object of classification. While Senft (2000: 36) raised a series of questions concerning what is classified, i.e., pure linguistic forms, or noun referents, Lucy (2000) provided a more definite criterion and argued that nominal classification is a classification of referents, for all forms of nominal classification contribute to "adequate noun phrase reference" (Lucy 2000: 329) and even the most grammaticalized gender in systems of agreement has referential functions (Lucy 2000: 330). Therefore, it can be concluded that no classification systems are realized by 'purely formal' or purely grammatical markers without indicating any semantics.

The definition of nominal classification given above allow us to distinguish between nominal classification systems and other means of classification. First, nominal classification devices are grammaticalized to some degree and occur in classificatory constructions. This feature differentiates nominal classification devices from such lexical means of classification as measure words and class terms, as measure terms are lexical terms with transparent semantics, while class terms are morphemes or words used as classificatory devices at word level (Grinevald 2000: 59-60; Senft 2007: 679). For example, the measure words in *a cup of water*, *a slice of bread*, and *a bundle of flowers* express quantity, while class terms are morphemes in compounds, e.g. *-berry* in *strawberry* and *blueberry* (Grinevald 2000: 59). Second, devices of nominal classification classify only

nouns and their referents. Those classification means involving events, actions and states are not regarded as nominal classification devices.

As shown in Table 1, nominal classification systems can be divided into different types, e.g., gender, classifiers, based on a variety of parameters, concerning such features as morphosyntactic realization, assignment, degree of grammaticalization, and semantic organization.

Table 1. Parameters of nominal classification (Aikhenvald 2000: 14-16)

morphosyntactic features	morphosyntactic loci or environments of nominal classification devices
scope, or domain of categorization	constructions or constituents that nominal classification devices occur in
assignment	semantic, morphological or phonological assignment
realization	affix, clitic, or independent words
agreement	presence or lack of agreement
markedness relation	whether or not they are functionally or formally marked
degree of grammaticalization	degree of obligatoriness
interaction with other grammatical categories	dependencies with such grammatical categories as number or case
semantic organization	based on universal or culture-specific parameters, the degree of semantic transparency
evolution and decay	source, development, and decay in use
language acquisition and dissolution	language acquisition (by children or adults) and dissolution (in aphasia)

In the following description, these parameters will be used to distinguish between gender and classifiers. Among these parameters, the degree of grammaticalization is one of the criteria often applied in typologies of nominal classification (cf. Aikhenvald 2000: 16-18; Grinevald 2000; Seifart 2010; Corbett and Fedden 2016). As shown in Fig. 1, gender and classifiers are placed on different positions on a continuum with regard to the degree of grammaticalization. Gender occurs on the right end as it is the most grammaticalized type, while classifiers appear in the middle as partly grammaticalized devices of nominal classification. As previously mentioned, lexical means with transparent semantics are not regarded as nominal classification systems.

<Lexical…………………………………….………….…………………grammatical>

measure/class terms   classifiers   gender

Fig. 1. Lexical and grammatical means of classification (adapted from Grinevald 2000: 61).

Other important parameters distinguishing gender and classifiers include morphosyntactic features, assignment, realization and agreement. With regard to morphological realization, gender can be marked on nouns, while classifiers are typically not affixed to nouns. In addition, gender systems differ from classifiers in terms of their assignment principles. Gender is assigned based on semantic principles or a combination of semantic and formal principles, while classifiers are more or less semantically motivated and they are assigned based on the properties of their noun referents. Agreement is regarded as the defining property of gender systems, in which gender is realized by distinctive forms in agreement within and/or outside the noun phrases. In contrast, classifiers are realized by morphosyntactic units marked only on one of the constituents (less likely on the noun itself) without agreement.

The above properties proposed in earlier typologies of nominal classification seem to show that there are clear boundaries between gender and classifiers. However, recent studies challenged such assumptions based on the evidence from languages involving intermediate and concurrent systems of gender and classifiers. Intermediate systems exist as a result of the grammaticalization of classifiers into gender systems and are characterized by properties that are traditionally ascribed to both gender and classifiers (Fedden and Corbett 2017). For example, in Ngan'gityemerri (or Nangikurrunggurr) (Southern Daly; Australia), 'classifier-like' classes shown in agreement are expressed by both free and bound forms, and the agreement markings are not as strictly obligatory as gender markings (Reid 1997). Concurrent systems of gender and classifiers have recently been documented in several studies. For example, in Mian (Trans New Guinea), there are four genders based on animacy and sex, and six verbal classifiers based on such semantics as sex, shape, and function (Corbett et al. 2017; Fedden and Corbett 2017).

Classifier systems can be further divided into different subtypes based on their morphosyntactic features. These include noun classifiers, numeral classifiers, genitive classifiers, verbal classifiers, and locative and deictic classifiers (Aikhenvald 2000). Noun classifiers occur with nouns, numeral classifiers appear adjacent to numerals in the context of quantification, genitive classifiers categorize nouns in possessive constructions, verbal classifiers are marked on verbs, and locative and deictic classifiers occur in locative and deictic noun phrases, respectively (Aikhenvald 2000: 1-4).

Systems of nominal classification "offer 'a unique window' into studying how humans construct representations of the world and encode them into their language"

(Aikhenvald 2000: 307). Nominal classification gives an insight into key phenomena such as agreement, the distinction between grammar and lexicon, and the functionality of grammatical categories. In addition, nominal classification systems are also regarded as a reflection of human cognition as well as cultural and social parameters. According to Aikhenvald (2000: 337-341), universal properties such as humanness and animacy for animate nouns and shape for inanimate nouns are the reflection of perceptual and cognitive mechanisms shared by humans, while other semantic parameters reflect culture-specific features, e.g., social status. Examples of such culture-specific categorization will be given in §2.4.2. In the sections §2.3 and §2.4, the two major types of nominal classification will be described in terms of their distribution, principles of assignment, morphosyntactic features and functions.

2.3. Gender

As mentioned in the previous section, gender is one of the two major types of nominal classification that are usually distinguished. In contrast with classifiers, gender systems are expressed by way of agreement, and are more grammaticalized. In the following overview, in §2.3.1, I will first describe the distribution of gender systems in the languages of the world, and then I will focus on assignment principles in §2.3.2, gender agreement in §2.3.3, and finally functions of gender will be reviewed in §2.3.4.

2.3.1. Distribution

Gender is widespread in languages in many parts of the world. As illustrated in the *World Atlas of Language Structures* (Corbett 2013a), gender occurs in 112 languages out of the 257 languages in the sample (see Fig. 2).
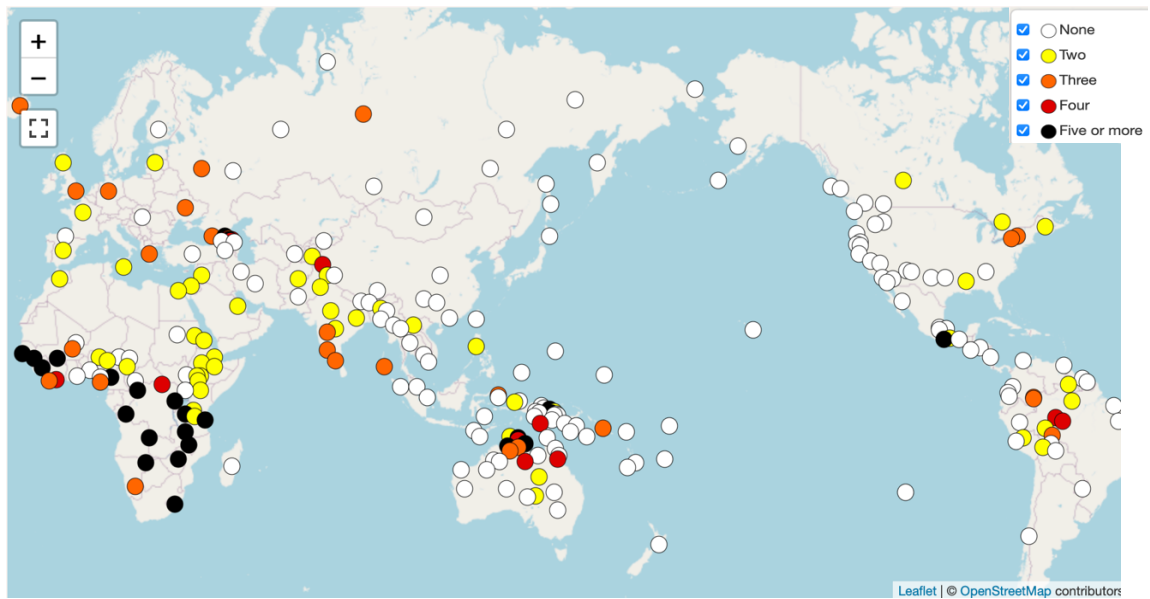
Fig. 2. Number of genders (Corbett 2013a).

The map shows the distribution of gender systems as well as the number of genders in languages with gender. Gender systems are predominantly found in the languages of Africa, e.g., in Niger-Congo and Afro-Asiatic languages, as well as most Indo-European languages spoken in Europe and South Asia. Gender also occurs in the languages of Australia and America. One of the areas where gender systems are not found is in East Asia, as will be shown in §2.4. With regard to the size of inventory, Corbett distinguishes among four types of systems, ranging from two to five or more genders. While smaller systems are found, e.g., in Indo-European and Afro-Asiatic languages, larger systems are found, e.g., in Niger-Congo languages and languages in Northern Australia.

2.3.2. Gender assignment
2.3.2.1. Semantic assignment

Nouns are assigned to genders based on two basic principles: semantic and formal. What should be noted is that all systems of gender are assigned based on a semantic principle, as there is a "semantic core" (Corbett 1991: 34) even in those gender systems based on formal principles. Most often, semantic principles involve sex as well as animacy and humanness. Other meanings which are relevant in semantic assignment include shape, and size (Aikhenvald 2017: 363).

12

As illustrated in the *World Atlas of Language Structures* (Corbett 2013b), in 53 languages out of 112 languages with gender, nouns are assigned based on semantic principles, while in 59 languages, nouns are assigned based on a combination of semantic and formal principles (see Fig. 3). Gender systems based on both semantic and formal assignment are found in Europe, South Asia, and Africa in Indo-European, Afro-Asiatic and Niger-Congo languages, while gender systems based on semantic assignment are scattered in South Asia, Australia, America, as well as other parts of the world.
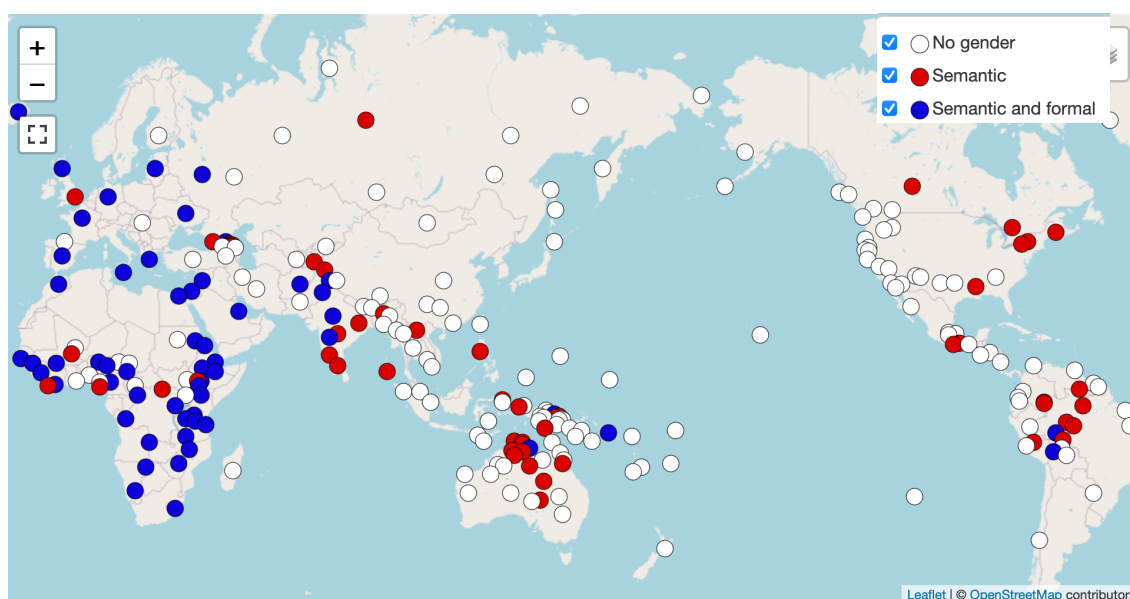
Fig. 3. Systems of gender assignment (Corbett 2013b).

In languages with only semantic assignment principles, semantic factors are thought to be 'sufficient' to account for the assignment of gender (Corbett 1991: 8). In such gender systems, two major types can be distinguished: strict semantic systems and predominantly semantic systems (Corbett 1991: 8-30).

In strict semantic systems, the gender of a noun can be inferred from its meaning and thus, genders are regarded as 'natural'. In Tamil, a Dravidian language spoken in south-east India and Sri Lanka and other parts of the world, there are two types of nouns: rational, i.e., masculine and feminine, and non-rational, i.e. neuter (Corbett 1991: 8-11). For example, nouns for gods and male humans are masculine, nouns for goddesses and female humans are feminine, and the residue is neuter.

Exceptions in assignment can occur in predominantly semantic systems. The fundamental principle of semantic assignment is still quite clear in these systems, although

there are residue classes. For example, in Dyirbal, a Pama-Nyungan language spoken in north-east Queensland, nouns are assigned to four genders (Dixon 1972: 308-312; see also Corbett 1991: 16-17). Gender 1 includes nouns for male humans and non-human animates, gender 2 includes nouns for female humans, water, fire, and fighting, gender 3 includes nouns denoting non-flesh food, and gender 4 is the residue comprising all those nouns not assigned to the former three. According to Dixon (1972), exceptions in Dyirbal can be explained in terms of semantic reassignment based on mythological association, concept association, and marking of important property. For example, instead of being assigned to gender 1, nouns for birds are assigned to gender 2, as birds are believed to be the spirits of dead human females, while *yarra* 'fishing line' is reassigned from gender 4 to gender 1 due to a conceptual association with fish (Corbett 1991: 17). However, Plaster and Polinsky (2007) proposed that gender assignment in Dyirbal is also motivated by formal features. For example, according to them, *yarra* 'fishing line' is assigned to gender 1 due to the similarity it shares with the word *yara* 'man' (Plaster and Polinsky 2007: 15-18). Such examples of reanalysis show that gender systems previously described as semantically based may in fact rely on formal properties as well.

In pronominal gender systems, nouns are assigned to genders based on the choice of anaphoric pronouns. An example is English, where gender is found in third person personal, possessive and reflexive pronouns. The choice of *he* and *she* is based on the natural gender of the referent, while the pronoun *it* is used to refer to inanimates. However, nouns in pronominal gender systems may be reassigned to a different gender. For example, ships and vehicles are frequently pronominalized as *she* (Wagner 2003: 1). According to Audring (2008: 107), the semantics of pronominal genders are aligned to a scale of "Individuation Hierarchy" as indicated in Fig. 4.

male human   > animal > inanimate object > mass/abstract female human
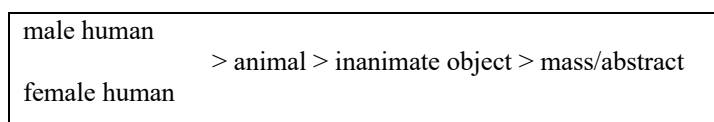
Fig. 4. Individuation Hierarchy (Audring 2008: 107).

By assigning nouns with a different pronoun, referents can be recategorized in terms of individuation hierarchy. For example, the personification of *wind* as *she* in Australian Vernacular English can be interpreted in terms of the individuation hierarchy (Pawley

2002: 159). *Wind* refers to an inanimate element of the natural physical environment and is uncountable. By being reassigned as feminine, *wind* moves from the least individuated class to the most individuated class with a differentiation of sexes.

In other languages, such as Russian (Slavic) and Swahili (Bantu), with gender systems based also on formal principles, semantic principles may still take precedence over formal principles based on the phonological shape or morphological structure of nouns. Such principles will be discussed in §2.3.2.2.

Unlike the gender languages mentioned above, Chinese does not have grammatical gender like many analytic or isolating languages of East and South-east Asia. However, it does have lexical morphemes that distinguish natural gender, e.g., *nán* 'male' and *nǚ* 'female'. In addition, analogies between pronouns and pronominal gender systems can also be identified in written Chinese. As shown in Table 2, there are three written forms of third person personal pronouns in modern Mandarin Chinese.

Table 2. Written forms of 3rd person pronouns in Mandarin Chinese

Person	Meanings	Singular	Plural
3rd	male	他 *tā*	他们 *tā-men*
	female	她 *tā*	她们 *tā-men*
	nonhuman	它 *tā*	它们 *tā-men*

These written forms are used to distinguish between masculine and feminine referents, as indicated in the opposition between 他 *tā* 'he' and 她 *tā* 'she', as well as humanness and animacy, as in 他 *tā* 'he' and 她 *tā* 'she' vs. 它 *tā* 'it'. It should be noted that the opposition is made only in written forms, with the pronouns pronounced in the same way.

In summary, nouns can be assigned to genders based on semantic rules. Semantic assignment principles are found both in gender systems based on solely semantic principles and in gender systems also involving formal assignment.

2.3.2.2. Formal assignment

Formal principles are another kind of criteria relevant in gender assignment. There are two types of formal principles: morphological and phonological. Morphologically, the gender of a noun can be assigned based on derivation or compounding, while phonologically, it can be distinguished based on the choice of sounds (Corbett 1991: 31). As

mentioned above, there are no purely formal systems: "formal assignment systems are really semantic plus formal systems" (Corbett 1991: 308).

Different assignment rules can either overlap or lead to a certain competition. Semantic and formal assignment principles may overlap in gender languages based on formal principles. For example, in German (Indo-European), neuter superordinate nouns occur with suffixes *Ge-*, as in *Getränk* 'beverage' (Rice 2006: 3-4). Gender assignment rules can also be shown in conflict. One example can be found in Russian, in which *djadja* 'uncle' and *deduška* 'grandfather' instead of being feminine like other nouns in declension II are masculine since they denote males (Corbett 1991: 38). Based on such examples, Corbett (1991: 68) proposes that "semantic factors usually take precedence" when there is competition between semantic and formal assignment rules. Another example can be found in German, in which superordinate nouns such as *Pflanze* 'plant' and *Waffe* 'weapon' are assigned to the feminine gender instead of neuter based on the suffix *-e*, while *Gemüse* 'vegetable' and *Gewerbe* 'trade, occupation' are assigned to the neuter gender based on their semantics as well as the initial '*Ge-*' instead of the final '*-e*' (Rice 2006: 5-6). Rice (2006) thus suggests that form and semantics contribute equally to gender assignment.

A noun's gender can be based on its inflection, typically in languages where every noun belongs to a morphological class. There is a substantial correlation between gender and the declension types in Russian (Slavic) and Swahili (Bantu) (Corbett 1991: 34-43). For example, in Russian, nouns in declension I, e.g., *zakon* 'law', are masculine, nouns in declensions II and III, e.g., *škola* 'school' and *kost'* 'bone', are feminine, and nouns in declension IV, e.g., *vino* 'wine', are neuter. As to indeclinable nouns, the gender of acronyms is determined by the gender of the head noun, e.g., *ŽÈK* (*žiliščno-èkspluatacionnaja kontora*) 'housing exploitation office' is feminine since the head noun *kontora* 'office' is feminine. The gender of other indeclinable nouns is determined by their semantics. For example, nouns denoting male humans, e.g. *attaše* 'attaché', are masculine, nouns denoting female humans, e.g. *ledi* 'lady', are feminine, nouns denoting animates, e.g. *kenguru* 'kangaroo', are masculine, while the residue, e.g. *taksi* 'taxi', is neuter (Corbett 1991: 40).

In addition to inflection, gender assignment can be based on derivation and compounding (Corbett 1991: 49-50). Derivational assignment can be illustrated by the role of suffixes and prefixes in German (Köpcke et al. 2010). For example, diminutives formed with *-lein* and *-chen* are neuter also in words denoting females, e.g., in *Mädchen* 'girl'

and *Fräulein* 'miss' (Köpcke et al. 2010: 173). In turn, the neuter gender of *Gesträuch* 'shrubbery' is determined by the collective prefix *Ge-* (cf. *Strauch* 'bush' masc.) (Zubin and Köpcke 1984: 45).

With regard to phonological assignment, a noun can be assigned a gender based on the choice of sounds, a sequence of sounds or suprasegmental features (Corbett 1991: 51-62; Aikhenvald 2000: 59-60). Gender assignment based on the choice of sounds occurs in Godié (Atlantic-Congo; Liberia and the Ivory Coast), which has three non-human genders assigned based on whether the final vowel of the noun stem is front, central or back (Corbett 1991: 53-54). Phonological rules can operate based on sequences of phonemes. As shown by Tucker et al. (1977), while most nouns in French ending in [ɔ̃] and [jɔ̃] are masculine, e.g. in *patron* 'boss', nouns ending in [ɛzɔ̃], [sjɔ̃], [zjɔ̃], [ɛjɔ̃], [tjɔ̃] are feminine, e.g. in *nation* 'nation'. Such gender assignment can be interpreted in terms of "backward processing" of sequences of penultimate and antepenultimate phonemes (Tucker et al. 1977: 62). Another example is provided by German, where complex consonant clusters occur in masculine monosyllabic nouns except in those nouns with clusters containing a non-sibilant fricative, i.e., [(C)+f, ç, x+t], which are predominantly feminine, as in *Luft* 'air' and *Frucht* 'fruit' (Köpcke and Zubin 1984: 29-32).

Gender can also be based on suprasegmental features involving tone and stress. For example, in Qafar, an East Cushitic language spoken in the Horn of Africa, the distinction between masculine and feminine can be made by shifting the position of accent: nouns ending in an accented vowel are feminine, e.g., *bariseynà* 'female teacher', while those with a non-final accent are masculine, e.g., *barisèyna* 'male teacher' (Corbett 1991: 51; Parker and Hayward 1985: 225).

In conclusion, the assignment of gender can be based on morphological and phonological principles. As mentioned previously, in such gender systems, semantic assignment is also relevant. As was shown in §2.2, the presence of formal agreement is one of the distinguishing features between gender and classifiers, because the choice of classifiers is not based on the shape of the noun.

2.3.3. Gender agreement

Agreement is regarded as the defining property of gender (Royen 1929: 526). It involves a "systematic covariance between a semantic or formal property of one element and a formal property of another" (Steele 1978: 610). Agreement is determined by the 'controller' ("the element which determines the agreement") and realized on at least one 'target' ("the element whose form is determined by agreement") (Corbett 2009: 342-343). It occurs in two main domains, as distinguished by Aikhenvald (2000: 29): the head-modifier domain between modifiers and heads within a noun phrase, and the predicate-argument domain between a predicate and its arguments within a clause. According to Greenberg's Universal 31, predicate-argument agreement presupposes agreement between adjectives and their head nouns in noun phrases (Greenberg 1963: 57). Apart from adjectives, agreement within noun phrases can also involve such modifiers as demonstratives, articles, possessives, numerals, and participles. As indicated in example (1) in Swahili (Bantu), the noun *ki-kapu* 'NCL7-basket' can be regarded as the controller, while the verb *ki-lianguka* 'NCL7-fell' can be regarded as an agreement target in the predicate-argument domain, which in turn presupposes the agreement with the adjective *ki-kubwa* 'NCL7-large' and the numeral *ki-moja* 'NCL7-one' as modifiers within the noun phrase.

(1) Alliterative concord in Swahili (Bantu) (Corbett 1991: 117; Welmers 1973)
  ki-kapu   *ki-kubwa*   *ki-moja*   *ki-lianguka*
  NCL7-basket NCL7-large NCL7-one NCL7-fell
  'One large basket fell.'

What should be noted is that agreement within the head-modifier domain also involves the agreement of relative pronoun and its head noun within a noun phrase, which may lead to agreement between personal pronoun and nouns outside the noun phrase and across clause boundaries. Gender agreement expressed on pronouns can be found in pronominal gender systems, such as in 3rd person singular pronouns in English. Such gender systems display "a minimum of formal exponence" by marking only on pronouns, and thus, are usually regarded as the least canonical type of agreement, or the 'extreme' systems of agreement (Audring 2008: 95).

 In opposition to the least canonical type of agreement, a canonical gender system is an idealized system of gender realized by consistent morphological markings on

controllers and targets across all domains within and outside the noun phrase. In such a canonical gender system, agreement is realized based on three criteria, quoted from Corbett and Fedden (2016: 505-520):

> Canonical gender-Criterion 1: Canonical gender values match agreement classes.
> Canonical gender-Criterion 2: In a canonical gender system the gender of a noun is constant across all domains in which a given language shows agreement.
> Canonical gender-Criterion 3: In a canonical gender assignment system, the gender of a noun can be read unambiguously off its lexical entry.

These criteria indicate that any violation of the three principles results in non-canonical agreement. In other words, a canonical gender system has not been attested, as there are exceptions involving, e.g., mismatches of gender values and agreement classes and inconsistencies of gender agreement in different domains, and cases where the gender of a noun is indicated by other sources of gender assignment than the lexical entry. In gender systems based on formal principles, the mismatches can still be identified between gender values and agreement classes. For example, in Swahili (Bantu), *Ø-baharia/ma-baharia* 'sailor/s' (Class 5-6) and *ki-nyozi/vi-nyozi* 'barber/s' (Class 7-8) take Class 1-2 prefixes on agreement forms instead of their respective classes based on their nominal prefixes (Contini-Morava 2002: 14).

Inconsistencies of gender agreement in different domains can be shown in the choice of gender agreement with hybrid nouns, which involve conflicts in gender assignment. As mentioned in §2.3.2.2, such nouns as *Mädchen* 'girl' and *Fräulein* 'miss' in German take feminine agreements in some contexts based on their meanings, but neuter in others based on the suffixes of *-chen* and *-lein* respectively. The choice of the agreement form with hybrid nouns is constrained by the Agreement Hierarchy, given in Fig. 5 below.

<div style="border:1px solid;padding:1em;text-align:center">
attributive > predicate > relative pronoun > personal pronoun
</div>
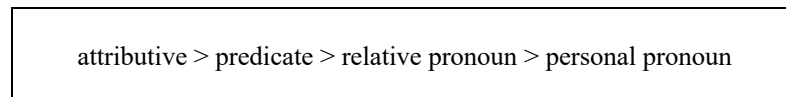
Fig. 5. The Agreement Hierarchy (Corbett 1991: 225-241; Corbett and Fedden 2016: 518-519).

The Agreement Hierarchy shows a rightward decrease in the choice of agreement form with hybrid nouns based on formal principles and a monotonical increase in the choice of agreement form based on semantic principles. In other words, semantic agreement with

hybrid nouns is possible in a position on the right part of the hierarchy, i.e., on a relative pronoun or a personal pronoun, typically when the distance between controller and target increases. This is illustrated in example (2) from German, in which the two pronouns *es* and *sie* show inconsistency in agreement. While the first pronoun is grammatically determined by the controller *das Mädchen*, the second pronoun occurs with an increasing distance from the controller.

(2) Grammatical and conceptual agreement in German (Köpcke et al. 2010: 190)
*Er fasste **das Mädchen** und zog **es** mit sich ins Wasser. Erschreckt und vor Angst schirie **sie** auf und rief um Hilfe.*
'He grabbed the girl [neut.] and pulled her [neut.] with him into the water. Terrified and frightened, she [fem.] screamed and shouted for help.'

While gender values of nouns can be read off the lexical entry in strict semantic systems, such as in Tamil (Dravidian), formal principles play a major role in gender assignment in many other languages. As shown in example (3) in French, compared with *homme* 'man' and *femme* 'woman', which take semantics as the only source of their gender assignment, many other nouns are assigned to a gender based on their forms rather than their semantics. As illustrated in the example, the suffixes *-eur* and *-rice* point to masculine and feminine genders respectively, based on which *intituteur* and *moteur* are masculine, while *institutrice* and *motrice* are feminine, regardless of whether the nouns are animate or inanimate.

(3) Gender in French
a.   masculine

le homme	'the man'
le instituteur	'the male schoolteacher'
le moteur	'the motor

b.   feminine

la femme	'the woman'
la institutrice	'the female schoolteacher
la motrice	'the power car'

All gender systems are shown in agreement realized in consistent markings on all or some elements related to nouns in a certain domain. This property is fundamental as to the distinction between gender and classifiers, as classifiers tend to be marked only on one of the elements within noun phrases. Furthermore, classifiers are less likely to occur on the noun itself, and therefore, their head nouns cannot function as controllers for agreement.

2.3.4. Functions of gender

Gender has both semantic and pragmatic functions. Contini-Morava and Kilarski (2013: 268-269) divide the semantic functions of gender into four subtypes: 1) "expansion of the lexicon", which involves the use of gender markers to create nouns; 2) "differentiating referents", by specifying properties of the referent by the choice of gender markers; 3) "individuation", by signalling information about number; 4) "ascribing properties to referents", involving information concerning attitudes toward the referents. As to the function of "expansion of the lexicon", gender markers typically found in morphological assignment systems can derive new lexical items. For example, in Swahili the noun *ki-zib*o 'plug' has been derived by attaching *ki-* (the prefix of NCL7) to the verb stem *-ziba* 'to plug up' (Contini-Morava 2002: 17-18).

Gender can also be used to differentiate noun referents by providing such information as sex, humanness and animacy, since all gender systems have a semantic 'core'. As indicated in (4) prefixed gender markers *r-* and *iak-* in Mohawk (Iroquoian) can be used to distinguish between male and female referents, in *rón:kwe* 'man' and *iakón:kwe* 'woman' (Mithun 2014: 156).

(4) Common nominals in Mohawk (Iroquoian) (Mithun 2014: 156)
    rón:kwe   *iakón:kwe*
    r-onkwe   iak-onkwe
    M.SG.AGT-be.a.person   FI.AGT-be.a.person
    'man'   'woman'

Thirdly, the choice of different gender markers may also express different degrees of individuation (Contini-Morava and Kilarski 2013: 276). For example, masculine and feminine gender markers *-o* (MASC. SG.) and *-a* (FEM. SG.) in Romance can be used to distinguish between count and mass nouns, as in *frutt-o* (count) 'fruit-MASC.SG.' and *frutt-a* (mass) 'fruit-FEM.SG.' in Italian, and *ram-o* (count) 'branch-MASC.SG. ' and *ram-a* (mass) 'branch-FEM.SG.' in Portuguese (Franco et al. 2015: 58).

Fourthly, gender can be used to express attitudes, e.g. affection or contempt, toward referents (Contini-Morava and Kilarski 2013: 277). For example, gender prefixes *vi-* (NCL8) and *ma-* (NCL6) in Swahili can be used to form diminutives and augmentatives. For example, assigning *vi-* to *toto* 'child' in the noun *vi-toto* 'little children' can express

a certain affection (Contini-Morava 2002: 17-18). In contrast, assigning *ma-* to *baba* 'father' in *ma-baba* 'AUG.-father' shows a degree of contempt (Contini-Morava and Kilarski 2013: 278), as indicated in example (5).

(5) Use of derived augmentatives to express contempt in Swahili (Bantu) (Contini-Morava and Kilarski 2013: 278)

mama	*mmoja*	*ma-baba*	*ishirini*
mother	one	AUG.PL-father	twenty

'one mother, twenty fathers'

As regards discourse functions, gender can be used for "referent identification" and "re-presentation of referents" (Contini-Morava and Kilarski 2013: 279-291). Firstly, gender markers in systems of agreement identify referents. In Nunggubuyu (Gunwinyguan; Northern Territory of Australia), genders are marked obligatorily on the verb, where affixes index referents in discourse. As is illustrated in example (6), prefixes denoting genders are found on all the three elements of subject, predicate and object, with both arguments marked on the verb. These gender markers make it possible to track referents and their case roles in discourse. More importantly, with these gender markers, referents can be identified even when the sentence is reduced to the inflected verb, as illustrated in (b).

(6) Anaphoric use of gender markers in Nunggubuyu (Gunwinyguan) (Heath 1983: 132)

a. *na-walyi-nʸunᵍ*   *nᵍu=na-nʸ*   *nᵍara-ma<u>n</u>i-nʸunᵍ*
MASC-male-HUM.SG   3MASC.SG/3FEM.SG=see-PAST   FEM-female-HUM.SG
'A/The man saw a/the woman.' (literally 'man he-her-see woman')

b. nᵍu= *na-nʸ*
3MASC.SG/3FEM.SG=see-PAST
'He saw her.'

With regard to "re-presentation of referents", the change of gender markers in discourse can indicate changes in perspectives on noun referents. In example (7) from a German translation of an interview with Alicia Silverstone in a teen girl's magazine, the German noun is first categorized by Silverstone as neuter as she follows the interviewer and assigns the related demonstrative, modifiers and relative pronoun to the neuter gender, and then the referent is recategorized by being referred to by feminine pronouns when Silverstone gives some positive comments on the portrayed character, as well as a result of the increasing distance of the pronouns from the controller, as described in example (2).

(7) Recategorization of referents in German (Zubin and Köpcke 2009: 241; cf.
     Contini-Morava and Kilarski 2013: 287)

Interviewe*r: In Ihrem neuen Kassenhit 'Clueless' spielen Sie **ein ziemlich schrilles
'material girl'**. Wie steht's da mit der Identifikation?*
Silverstone: **Das** *ist* **ein sehr oberflächliches Mädchen, das** *nur an Klamotten und
Geld denkt – als Rollenvorbild für Kids natürlich denkbar ungeeignet. Andererseits
ist* **sie** *sehr modern, sehr zeitgemäß – es war schon lustig, sie zu spielen. Interview
mit Alicia Silverstone, 18 Jahre. Allegra 11/95, 189.*
Interviewer: *In your new box-office hit 'Clueless' you play* **a rather shrill 'material
girl'** *(neut.). Is there any identification going on there?*
Silverstone: **That** *(neut.) is* **a very superficial girl** *(neut.), who (neut.) thinks only
about clothing and money – not particularly appropriate as a role model for kids.
But otherwise* **she's** *(fem.) very modern, very contemporary – it was really fun to
play* **her** *(fem.).*

The above functions show that gender has all the semantic and pragmatic functions shared
by the systems of nominal classification. Gender markers can be used to differentiate
referents and show different degrees of individuation and attitudes toward referents, and
to identify and recategorize referents in discourse. In contrast with classifiers, gender is
more characteristic of the "expansion of the lexicon", while less typical of "reference
management" in discourse. The function concerning "reference management" in dis-
course will be further analysed in §3.3.2 in the context of the functions of classifiers.

2.4. Classifiers

Classifiers are widespread as devices of nominal classification. They are defined as mor-
phemes occurring "in surface structures under specifiable conditions" to denote "some
salient perceived or imputed characteristic of the entity to which an associated noun refers"
(Allan 1977: 285). Classifiers can be divided into numeral classifiers, noun classifiers,
verbal classifiers, genitive classifiers, locative and deictic classifiers. In the following
sections, these types will be reviewed in terms of their distribution, semantics, and mor-
phosyntactic features.

2.4.1. Types of classifiers

2.4.1.1. Numeral classifiers

Numeral classifiers are one of the most well-known nominal classification devices. They always occur within numeral noun phrases and classify noun referents in terms of their inherent properties. The distribution of numeral classifiers and gender is largely complementary, as shown in

Fig. 6. While gender is predominant in most Indo-European and African languages, numeral classifiers are nearly absent in the languages in these regions, particularly in Europe. According to Gil (2013), only one language in Europe, Hungarian, has the optional use of numeral classifiers. Numeral classifiers occur mostly in languages in East and Southeast Asia, including Chinese, with the rest scattered in other parts of the world, e.g., South Asia and America. Therefore, numeral classifiers are one of the most characteristic features differentiating languages in East and Southeast Asia from Indo-European languages, or more specifically differentiating Chinese from English.
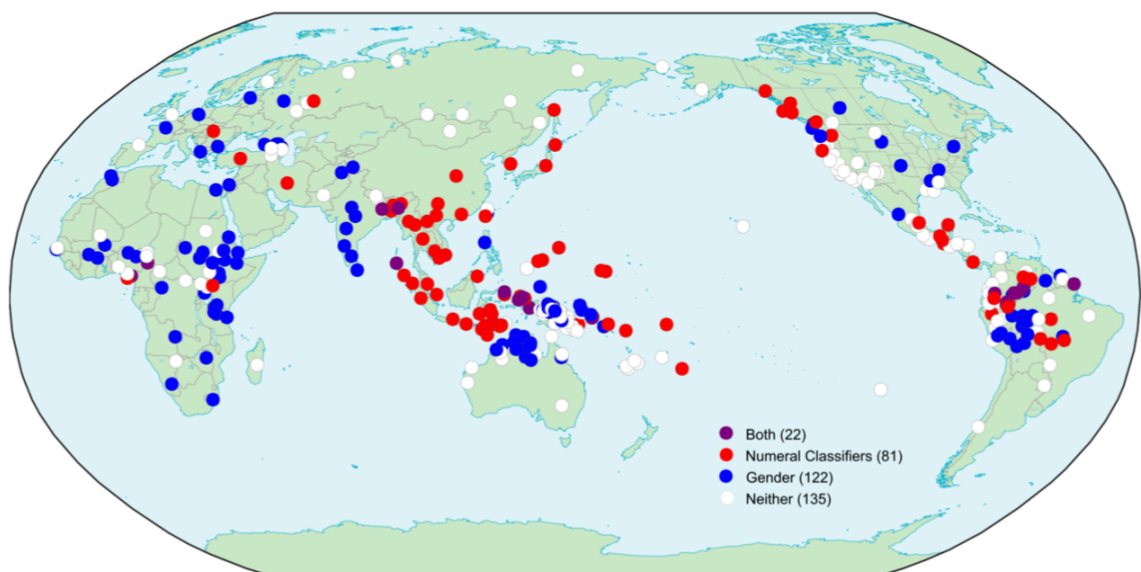
Fig. 6. Distribution of numeral classifiers and gender (Sinnemäki 2019: 151).

Several terms have been used to refer to this type of classifiers: classifiers (e.g., Tai and Wang 1990; Cheng and Sybesma 1998; Simpson 2005), numerative classifiers (Becker 1975, 1986), noun classifiers (Sanches and Slobin 1973; Erbaugh 1986), and nominal classifiers (T'sou 1976). In Chinese linguistics, they tend to be confused with quantifiers, as they occur in the same slot in numeral noun phrases.

Numeral classifiers can be independent words occurring next to numerals or quantifiers typically in isolating languages, and less frequently in agglutinating and fusional languages (Aikhenvald 2000: 101-103). As shown in example (8) the numeral classifier *orang* 'CLF:ANIMATE:HUMAN' in Malay, an Austronesian language, is an independent word occurring adjacent to the numeral *tiga* 'three'.

(8) Numeral classifiers in Malay (Salehuddin et al. 2011)
tiga	*orang*	*kanak-kanak*
three	CLF:ANIMATE.HUMAN	child
'three children'

Numeral classifiers can also be affixed to numerals as suffixes, or less frequently as prefixes. Such numeral classifiers are most common in the languages of Americas and in inflecting Indic languages (Aikhenvald 2000: 105), and they can also be found in some languages in South, Southeast and East Asia, e.g. Japanese. As indicated in (9) the numeral classifier *-dai* 'CLF:VEHICLE' in Japanese is a suffix attached with the linker particle *-no* to the numeral *ni-* 'two'. Numeral classifiers can also be clitics fused with a numeral, typically in languages with fusional characteristics.

(9) Numeral classifiers in Japanese (Aikhenvald 2000: 106)
ni-dai-no	*kuruma*	*o*	*kai-mashi-ta*
two-CLF:VEHICLE-LINKER	car	ACC	buy-HON-PAST
'(S/he) bought two cars.'

The size of numeral classifier systems varies in different languages. In some languages, there is only one or a small number of numeral classifiers. For example, Nung (Tai, Vietnam, China and Laos), has only four numeral classifiers (Aikhenvald 2000: 103). In contrast, a large classifier system may have a dozen or even hundreds of numeral classifiers. For example, there are at least 154 numeral classifiers in Korean (Lee 2014), and approximately 150 in Japanese (Yamamoto and Keil 2000). However regardless of the size of the classifier system, there is usually one general classifier, e.g., *hon* in Japanese, which can be applied to most nouns in the languages concerned.

The choice of a numeral classifier is always semantically motivated. They tend to express meanings concerning humanness and animacy. Based on these classifications, numeral classifiers can further categorize humans in terms of their social function and status, and inanimates based on such physical properties as shape, dimensionality,

extendedness, directionality, and interioricity (Aikhenvald 2000: 286-291). For example, in Korean, the numeral classifier *myengi* 'CLF:HUMAN' has a neutral meaning, as in *gyosu* 'professor' in *gyosu ne myengi* 'four professors' (Kim 2005: 219). In contrast, *nom* 'CLF:DESPICABLE OR INSIGNIFICANT MALE' and *nyen* 'CLF:DESPICABLE OR INSIGNIFICANT FEMALE' refer to humans of lower social status, as shown in example (10).

(10) Korean numeral classifiers referring to humans of lower social status
(Lee 2014: 42-43)
    a.   *kkangphay*   *twu*   *nom*
        gangster   two   CLF:DESPICABLE OR INSIGNIFICANT MALE
        'two male gangsters'
    b.   *kkangphay*   *twu*   *nyen*
        gangster   two   CLF:DESPICABLE OR INSIGNIFICANT FEMALE
        'two female gangsters'

Other Korean numeral classifiers for animates and inanimates include *mali* 'CLF:ANIMAL' and *calwu* 'CLF:LONG'. As shown in example (11), *mali* 'CLF:ANIMAL' refers to animacy, while *calwu* 'CLF:LONG' indicates the shape of the referent of the inanimate noun *yenphil* 'pencil'. What should be noted is that there are several possible orders of numeral noun phrases in Korean, as long as numeral classifiers occur with numerals in the same constituent (cf. Aikhenvald 2000: 106).

(11) Korean numeral classifiers denoting animacy and inanimacy (Lee 2014: 23)
    a.   *sey*   *mali*   *kom*
        three   CLF:ANIMAL   bear
        'three bears'
    b.   *sey*   *calwu*   *yenphil*
        three   CLF:LONG   pencil
        'three long pencils'

Numeral classifiers can be distinguished from measure terms or quantifiers in nonclassifier languages based on such properties as applicability, semantics, pragmatics and the degree of grammaticalization as indicated in Table 3.

Table 3. Differences between numeral classifiers and measure words

(cf. Dixon 1982a; 1982b: 213-218; 1986)[3]

Property	Numeral classifiers	Measure words
applicability	limited groups of nouns	wider variety of nouns
semantics	qualities	quantities
pragmatics	tracking the referents	not tracking the referents
grammaticalization	partly grammaticalized	not grammaticalized

According to Ahrens (1994: 204), numeral classifiers can only be applied to "a limited and specific group of nouns" , while measure words tend to be used as "a measure for a wide variety of nouns". This criterium shows the semantic correlation of numeral classifiers and their head nouns. However, it is not predictive enough to distinguish between numeral classifiers and measure words, as some measure words are also applied to a restricted set of nouns. For example, *flock* in English is only used with nouns referring to sheep or birds of the same type, while *herd* cooccurs with nouns referring to cows, deer, or elephants. A fundamental property to differentiate numeral classifiers from measure words lies in their semantics. Numeral classifiers denote the "essential" properties of noun referents, while measure words or quantifiers provide "accidental" information regarding quantity (Aikhenvald 2000: 117; Her 2012: 10). The properties that numeral classifiers specify are thus inherent in their referents. However, the information provided by measure words is irrelevant to the nature of noun referents. Such semantic properties are also shown in the use of numeral classifiers and measure words in discourse. Numeral classifiers, encoded with inherent features of noun referents, can be used to refer to their referents in the context, while measure words cannot be used in the same way without the intrinsic association with noun referents (Aikhenvald 2000: 118). In terms of the degree of grammaticalization, numeral classifiers are partly grammaticalized and are used to fill an obligatory slot in numeral noun phrases. In contrast, measure words are not grammaticalized and can cooccur relatively more freely with nouns (see Grinevald 2000; Aikhenvald 2000: 114-120). However, in classifier languages, e.g., Chinese, with both numeral classifiers and measure words, measure words are also grammaticalized, as they are obligatory and occur in the same slot of numeral classifiers in numeral noun phrases, although they can be used with a wider variety of nouns.

In classifier languages, two subtypes of numeral classifiers can be distinguished: sortal classifiers and mensural classifiers (Aikhenvald 2000: 118; Her and Hsieh 2010).

---

[3] See also Ahrens (1994), Aikhenvald (2000: 286-291), Grinevald (2000: 61), Her (2012).

Sortal classifiers are 'true' numeral classifiers and are usually omitted when translated into non-classifier languages, as indicated in example (11) above. In contrast, mensural classifiers, occurring in the same structure as sortal classifiers, are similar to measure terms, and thus, they have equivalents in non-classifier languages. In English, the equivalents of mensural classifiers tend to occur in a 'pseudopartitive' structure of [NUM+M+of+N]. [4] As shown in example (12), the numeral classifier *twulum* 'CLF:STRING" in Korean is equivalent to the partitive noun *string* in English. However, partitive nouns in such structures in English should be regarded as measure words rather than numeral classifiers, since they are not grammaticalized and none of them can be regarded as general enough to be applied to most nouns, as general classifiers do in classifier languages.

(12) Mensural classifiers in Korean (Lee 2014: 73)
> *chenge*   *twu*   *twulum*
> herring   two   MENS:STRING
> 'two strings of herring'

In contrast, Senft (2007) argues that there are no grounds for the distinction between sortal and mensural classifiers, as the distinction between numeral classifiers and measure words is sufficient for the classification of different types of classifiers. In this dissertation, I will use the terms 'numeral classifiers' and 'measure words' as terms for the two kinds of elements occurring within numeral noun phrases in classifier languages.

In summary, numeral classifiers are one of the characteristic features that distinguish the languages of East and Southeast Asia from Indo-European languages. Numeral classifiers occur with numerals as independent words, affixes or clitics, and tend to denote humanness, animacy and physical features of nominal referents.

2.4.1.2. Noun classifiers

Noun classifiers occur in noun phrases but are 'independent' of any other constituents within or outside the noun phrase (Aikhenvald 2017: 367). They can be found in

---

[4] For discussions about the properties of 'pseudopartitive' structure in generative accounts, see Jackendoff (1977), Selkirk (1977), Beckwith (2007: 38-66), Alexiadou et al. (2007: 395-438).

Australian, Mesoamerican, Western Austronesian and Oceanic languages. Noun classifiers also occur in Tai, Tibetan and some Austroasiatic languages. They are sometimes referred to as 'generic classifiers' or 'generics' especially in the Australianist tradition (Harvey and Reid 1997: 9-10; Sands 1995: 269-270; cf. Aikhenvald 2000: 81). The size of noun classifier systems varies from just two in Emmi (Western Daly; Australia) (Ford 1998) to twenty in Yidiny (Pama-Nyungan; Australia) (Dixon 1982c). In some extreme cases, almost all generic nouns can be used as noun classifiers, as in Minangkabau, an Austronesian language spoken in Indonesia (Marnita 1996; cf. Aikhenvald 2000: 189).

There are several differences between noun classifiers and numeral classifiers. Noun classifiers are affixes marked on nouns or independent lexemes adjacent only to nouns. Furthermore, noun classifiers generally do not occur in the context of quantification, except in noun phrases in which there are numeral classifiers marked on numerals and noun classifiers attached to nouns as shown in example (13) in Ersu, a Tibeto-Burman language from China.

(13) Noun classifiers in Ersu (Zhang 2013: 314)

pi=kaka	*si-wo*	
potato = CLF:ROUND AND LARGER THAN A FIST	three-CLF:GENERIC,	NON-STICK-LIKE

'three potatoes'

Although noun classifiers can be affixes attached to nouns, as shown in example (13), they are usually expressed through independent lexemes. In example (14) in Jacaltec, a Mayan language of Guatemala, *naj* 'CLF:MAN/MALE' and *no7* 'CLF:ANIMAL' are independent lexemes that cooccur with nouns and indicate inherent properties of their referents.

(14) Noun classifiers in Jacaltec (Craig 1986a: 264)

a.
xil	*naj*	*xuwan*	*no7*	*lab'a*
saw	CLF:MAN	John	CLF:ANIMAL	snake

'John saw the snake'

b.
xil	*naj*	*no7*
saw	CLF:MALE	CLF:ANIMAL

'He saw it'

Noun classifiers are akin to lexical terms in non-classifier languages, e.g. compound elements such as *-berry* in *strawberry* and *-man* in *policeman* (Grinevald 2000: 59-61; Aikhenvald 2017: 370). However, similar to the distinction between numeral classifiers and quantifiers, noun classifiers are more closely related to nouns based on their inherent features and are more grammaticalized.

The choice of noun classifiers is based on such meanings as sex, animacy, social status for animates, and material and function for inanimates. For example, Yidiny has twenty noun classifiers, which can be categorized into two semantic groups: inherent nature classifiers for humans, animals and inanimates, and function classifiers such as edible, habitable, drinkable, etc. (Dixon 1982c: 192-199; cf. Aikhenvald 2000: 83). Apart from sex-based noun classifiers, in Jacaltec humans are further classified into different classes according to their kinship and age, while inanimates are further categorized based on material (Craig 1986b: 266-284).

In conclusion, noun classifiers are marked on nouns or occur with nouns as independent words. They are used to categorize noun referents based on such inherent properties as humanness, animacy and material.

2.4.1.3. Verbal classifiers

Verbal classifiers appear on the verb and classify the referent of its argument, in subject function (for intransitive verbs) and object function (for transitive verbs). They are not as common as the two types of classifiers discussed above and are mainly found in North America, South America and Northern Australia. Verbal classifiers are not attested in the languages of Africa and Eurasia and in the Austronesian family (Aikhenvald 2000: 169-171). The inventories of verbal classifiers vary from two, such as in Tibeto-Burman languages (LaPolla 1994: 75), to as many as 100, as in Imonda, a Papuan language from the Waris language family (Seiler 1986).

According to Aikhenvald (2000: 150-162), verbal classifiers can be divided into three subtypes: incorporated verbal classifiers, classifying affixes and suppletive classificatory verbs. Incorporated verbal classifiers have a generic-specific relationship with the referent, i.e., they are usually more generic and function as superordinate terms for the noun, either the subject or object. The other two subtypes of verbal classifiers tend to

specify inherent properties of the referent in terms of animacy, shape, consistency, arrangement, and orientation in space.

Incorporated verbal classifiers are also referred to as incorporated generic nouns (Aikhenvald 2000: 150-162), as they are realized through noun incorporation into the verb, with the generic nouns still recognizable. In example (15) in Cayuga (Northern Iroquoian; Ontario), the incorporated nouns *náhskw* 'CLF:DOMESTIC ANIMAL' and *'tręht* 'CLF:VEHICLE' are superordinate terms for *so:wá:s* 'dog' and *k'atrehta'* 'car' respectively.

(15) Incorporated verbal classifiers in Cayuga (Northern Iroquoian) (Mithun 1986: 387-388)
   a. *So:wá:s*   *akh-náhskw-aę'.*
      dog   I-CLF:DOMESTIC ANIMAL-have
      'I have a (pet) dog.'
   b. *K'atrehta'*   *ake-'tręht-áę'.*
      car   I-CLF:VEHICLE-have
      'I have a car.'

The second subtype of verbal classifiers are phonologically eroded affixes occurring on verbs. These classificatory verbal affixes classify referents based on physical properties, such as shape, dimensionality, size, interiority, consistency, constitution and arrangement. As illustrated in example (16) from Innu, an Algonquian language spoken in north-eastern Quebec and Labrador in Canada, the affix *-êk-* 'CLF:SHEETLIKE' classifies the subject *tâpiškâkan* 'scarf', based on its shape, while, *-pêc-* 'CLF:FLOWING.LIQUID' specifies the nature of water.

(16) Classifying affixes in Innu (Drapeau and Lambert-Brétière 2011: 302-303)
   a. *mâk*   *ni-tâpiškâkan*   *mišta-miš-eci-ší-pan*
      and   1-scarf   very-big- CLF:SHEETLIKE-AI-PRT
      'And my scarf was very large.'
   b. *nê*   *wâpikunâpu-lu*   *nûtim*   *ni-šišu-pêc-in-âw*
      DEM   Florida.water-OBV   all   1-rub-CLF:FLOWING.LIQUID
                                       by.hand.TA-3.3'
     anitê   *u-yâ-t,*   *êkwê*   *aštêpitiku-t*
      on   3-body-LOC   then   stop.cramps.AI-3CJ
      'I rub her completely with Florida water and then she stopped cramping.'

Finally, in classificatory verbs, the choice of a verb stem is determined by inherent properties of concrete objects in motion or at rest, such as shape, form, animacy, number, and consistency. LaPolla and Huang (2003) refer to classificatory verbs as existential or

locative verbs. According to them, Qiang, a Tibeto-Burman language, has five existential/locative verbs based on animacy and location (LaPolla and Huang 2003: 133-134). As indicated in example (17), the distinction between animate and inanimate referents in this language is made by *ẓi* for the former, and *ṣə* and *xu* for the latter. All of the five existential classificatory verbs can also indicate the properties concerning the location of the referent: *ṣə* and *xu* are for referents "not in containers or immovable or inalienably connected to some larger entity", *le* is for referents "located in a containment of some type", and *we* is for "immovable referents, referents inalienably connected to a larger entity, and happenings" (LaPolla and Huang 2003: 133).

(17) Existential and locative verbs in Qiang (LaPolla and Huang 2003: 133-134)
   a. *tʂuats-məq-ta*   *ləɣz-e-pen*   *ṣə.*   (inanimate)
      table-top-LOC   book-one- CLF   exist
      'There is a book on the table.'
   b. *ku-ta*   *pi-a-la*   *xu.*   (inanimate)
      floor-LOC   pen-one-CLF   exist
      'There is a pen on the floor.'
   c. *zə-ʁa*   *jajŋu̩*   *le.*   (inside vessel)
      ground-LOC   potato   exist
      'There are potatoes in the ground.'
   d. *səf-tho-ʐgu-ta*   *wətshi-o-u*   *-ʐi.*   (animate)
      tree-that:one-CLF-LOC   sparrow-one- CL   exist
      'There is a sparrow in that tree.'
   e. *qəl-la*   *səf-o-ʐgu*   *we.*   (immovable, connected to
      below-LOC   tree-one-CLF   exist   major entity)
      'There is a tree below.'
   f. *qu*   *dzoqu-ji-tua*   *wa.*   (immovable, connected to
      1SG   leg-two-CLF   exist:1SG   major entity)
      'I have two legs.'
   g. *the:*   *tʂheχuo*   *de-w(e).*   (happening)
      3SG   car.accident   DIR-exist
      'I have a car accident.'

The choice of a classificatory stem may also be based on orientation in space. As shown in Table 4, in Nevome (Uto-Aztecan; Arizona and Mexico) there are four pairs of classificatory verbs used to classify the referent of the verb's argument as standing or lying, and animate or inanimate.

Table 4. Classificatory verbs in Nevome (Shaul 1986: 12)

	Inanimate referent	Animate referent
be lying	*catu/vutu*	*voho/vopo*
be standing	*cuhca/tutu*	*cuhca/guguhuca*

Classificatory verbs are analogous to lexical classification in such non-classifier languages as English, where the choice of a verb depends on the properties of the object in 'semantic agreement', e.g., in such verbs of such as *suck* (a hard object), *drink* (a liquid), *chew* (a denser object). Therefore, Grinevald (2000: 68) excluded classificatory verbs from the subtypes of verbal classifiers as a lexical classification ubiquitous in any language. However, according to Aikhenvald (2000: 153), classifier languages possess distinguishable paradigmatic sets among verbs for handling objects based on their physical properties or positions, and Allan (1977: 289-290) also pointed out that classifier languages also assign some formal markers to classificatory verbs as exemplified above. In contrast, in non-classifier languages, semantic agreement of verbs and objects occurs in a very small number of verbs, such as verbs of consumption in English. In addition, there is no formal marking on verbs in languages such as English.

2.4.1.4. Genitive classifiers

Genitive classifiers occur in possessive noun phrases. They are labelled as 'classifiers in possessive constructions' by Aikhenvald (2017: 375), and as 'attributive', 'possessive' and 'relational' classifiers by Grinevald (2000: 66). They classify limited sets of nouns based on the properties of the possessed noun referents, the relation between the possessee and the possessed, and the properties of the possessor, based on which they are divided into three subtypes as possessed classifiers, relational classifiers, and possessor classifiers (Aikhenvald 2000: 127-147). Genitive classifiers can be realized as independent words, affixes or clitics attached to nouns or genitive markers, or clitics attached to possessor nouns.

According to Aikhenvald (2000: 127-147), genitive classifiers are rarer compared with noun classifiers and numeral classifiers. They are absent in Australian and Eurasian languages. Possessed classifiers are found in North and South American languages, and one Niger-Congo language, Dongo-ko. They can also coexist with other types of

classifiers, e.g., numeral classifiers, in such languages as Hmong and other Miao-Yao languages spoken in Northern China and Indochina, and Papuan languages of Central and Southern Bougainville. Relational classifiers tend to coexist with numeral classifiers in Oceanic and Micronesian languages. However, only two South American languages have relational classifiers. Possessor classifiers are extremely rare and have been attested only in Makú languages in north-western Amazon, such as Dáw.

The choice of genitive classifiers is always semantically motivated. Possessed classifiers categorize the referent of the possessed noun in terms of such properties as animacy, shape, size and function. For example, in Hualapai (Yuman; Arizona), *-hat* is used to refer to a pet or domestic animal, while *-wi:nych* is used as a general classifier (Aikhenvald 2017: 376). Relational classifiers indicate a possessive relation between nouns. As indicated in example (18), the general relational classifier in Boumaa Fijian, an Austronesian language, indicates the possessor of the gun.

(18) General relational classifiers in Boumaa Fijian (Dixon 1988: 137)

a	*o-mu*	*da'ai*
ART	CLF-2SG	gun

‘your gun’ (which belongs to you)

Relational classifiers can also specify the way in which possessors handle possessed objects as well as their value. Such uses are illustrated in example (19) from Kipeá, an extinct Karirian language of Brazil.

(19) Use of relational classifiers to indicate different ways of handling possessum in Kipeá-Kariri (Rodrigues 2012: 261)

a.
dz-upodó	*do*	*buké*
1SG- CLF:ROASTED	POSS	deer

‘my deer (roasted)’

b.
dz-udé	*do*	*ghinhé*
1SG-CLF:GATHERED.IN.THE.GARDEN	POSS	beans

‘my beans (from my garden)’

Compared with other types of classifiers, relational classifiers more often denote the value of an object to a possessor. As indicated in example (20) in Raga, an East Vanuatu language from the Austronesian family, *pila-* ‘CLF:VALUABLE.POSSESSION’ indicates the value of *que* ‘pig’ to the possessor.

(20) Use of relational classifier to show the value of a possessum in Raga
     (Lichtenberk 1983: 154)

 que   *pila-ma*
 pig   CLF:VALUABLE.POSSESSION-2SG
 'your valued pig'

Finally, possessor classifiers distinguish between animate and inanimate possessors. For example, in Dâw, a Nadahup language spoken in the north-western part of Brazil, there are two possessor classifiers in the form of clitics: *-ẽj* and *-dee'* (Martins 1994: 138-141; cf. Aikhenvald 2000: 139).

(21) Possessor classifiers in Dâw (Martins 1994: 138-141)
 a. *yud* *dâw*   *tôg-ẽj*
   cloth- human   daughter-CLF:ANIMATE.POSSESSOR
   ing
   'The clothing is girl's; the girl's clothing'
 b. *yud* *kaw-wá'-dee'*
   man- garden-up- CLF:INANIMATE.POSSESSOR
   ioc
   'manioc of a garden'

As indicated in example (21), *-ẽj* denotes an animate possessor in (a), while *-dee'* denotes an inanimate possessor in (b).

2.4.1.5. Locative and deictic classifiers

Locative and deictic classifiers are not very common in the languages of the world, so they are usually referred to as 'minor' classifiers in contrast with the above four types (Grinevald 2000: 63, 68-69). They have only been found in the languages of North and South America, and their inventories are usually relatively smaller than those of the other types of classifiers.

 Locative classifiers appear on adpositions in locative noun phrases, and denote the physical properties, e.g., shape or consistency, of the argument of the adposition. For example, in Lokono (North Arawak), there is a small system of five locative classifiers which indicate properties of the referent, as in *loko* 'inside a hollow or solid object', *rakon* 'in a fluid', and *kolokon* 'in fire or light', and interiority and dimensionality as in *roko* 'on the inside surface of' (Pet 2011: 20-21). Locative classifiers are comparable with

prepositions in such non-classifier languages as English in that the choice of a locative classifier depends on the properties of the referent of the noun (Aikhenvald 2000: 176).

Deictic classifiers are obligatory elements attached to or fused with articles and demonstratives, and occur in, e.g., Siouan and Eskimo languages as well as Yuchi in North America, and Guaicuruan languages in South America. Their inventories are small. For example, three to twelve deictic classifiers are found in Siouan languages (Rankin 2004).

Deictic classifiers in Siouan and Guaicuruan languages developed from grammaticalized verbs (Aikhenvald 2000: 182). Deictic classifiers in Siouan languages are used to denote stance (for animates) and shape for inanimates (Rankin 2004). For example, *ðįkhe* 'sitting', *thǫ* 'standing', *khe 'lying'*, and *ðį* 'moving' can be attached to nouns referring to animates (Rankin 2004: 211-212). In Guaicuruan languages, deictic classifiers are used to denote properties of their referents in terms of spatial position/location, extendedness, extension and visibility.

2.4.2. Functions of classifiers

While classifiers share all the main types of functions with gender, there are some differences resulting from the different ways in which they are expressed. As in §2.3.4, two types of functions will be reviewed, i.e., semantic and discourse functions.

As for semantic functions, the function of lexical expansion is not very common in classifier systems. Classifiers affixed to nouns are more likely to cooccur with other nominal classification devices, e.g. class terms in Lao (Tai), or in multiple classifier systems, e.g. Mopan Maya, which has both numeral classifiers and noun classifiers (Contini-Morava and Danziger 2018; Danziger and Contini-Morava 2020). In these systems, class terms or noun classifiers similar to gender markers in terms of morphological features are used to derive new nouns from noun or verb stems. For example, the class term *mè^0* (a phonologically reduced form of the noun *mèè^1* 'mother') in Lao can be used to derive new nouns to indicate female occupations, e.g. in *mè^0-khaaw^3* 'nun' (cf. *khaaw^3* 'white') or *mè^0-caang^4* 'prostitute' (cf. *caang^4* 'hire someone's services') (Enfield 2004: 136).

Classifiers in general are more frequently used to differentiate and individuate nouns and to ascribe properties to noun referents (Contini-Morava and Kilarski 2013:

268-279). Firstly, classifiers can differentiate noun referents and express different properties when the same noun is used with different classifiers. For example, numeral classifiers in Burmese can be used to show different properties of noun referents. As indicated in Table 5, *myiʔ* 'river' in Burmese can be used in at least eight contexts with different numeral classifiers, resulting in different interpretations of the noun.

Table 5. Reclassification of an inanimate noun in Burmese (Becker 1975: 113; cf. Aikhenvald 2000: 319)

Noun	Numeral	Classifier	Translation
myiʔ	*tə*	*yaʔ*	'river one place' (e.g., destination for a picnic)
myiʔ	*tə*	*tan*	'river one line' (e.g., on a map)
myiʔ	*tə*	*hmwa*	'river one section' (e.g., a fishing area)
myiʔ	*tə*	*'sin*	'river one distant arc' (e.g., a path to the sea)
myiʔ	*tə*	*θwɛ*	'river one connection' (e.g., tying two villages)
myiʔ	*tə*	*'pa*	'river one sacred object' (e.g., in mythology)
myiʔ	*btə*	*khu'*	'river one conceptual unit' (e.g., in a discussion of rivers in general)
myiʔ	*tə*	*myiʔ*	'river one river' (the unmarked case)

All types of classifiers can thus be used to specify noun referents. However, there is one exception in classifier systems regarding semantic specificity. Incorporated verbal classifiers are used to denote a more generic domain of the referents rather than to specify their properties, as discussed in §2.4.1.3.

Secondly, classifiers, in particular numeral classifiers, can be used to individuate nouns. Numeral classifiers are usually regarded as complementary to plural marking systems (Borer 2005: 92-96). According to Borer (2005: 93), while in some languages, e.g., English, nouns are individuated by plural inflection and indefinite articles, in classifier languages nouns are individuated by numeral classifiers. Borer (2005: 95) further points out that in such languages as Armenian (Indo-European), plural markers and numeral classifiers do not cooccur, since plural markers occur as suffixes and numeral classifiers are independent pre-nominal morphemes. As shown in example (22), the plural *-ner* in (a) is marked on the noun *hovanoc* 'umbrella' without the numeral classifier *had,* which occurs with the noun in (b).

(22) Mutual exclusiveness of numeral classifiers and plural markers in Armenian
 (Indo-European) (Borer 2005: 94-95)
 a. *yergu*   *hovanoc-ner*   *uni-m.*
    two   umbrella-PL   have-1SG
    'I have two umbrellas.'
 b. *Yergu*   *had*   *hovanoc*   *uni-m.*
    two   CLF   umbrella   have-1SG
    'I have two umbrellas.'

Numeral classifiers are obligatory in classifier languages, because nouns in classifier languages are regarded as concept nouns, similar to mass nouns in English, and should be individuated before they are combined with numerals (Lucy 1992: 73). Such arguments were corroborated by Chierchia (1998a, 1998b), who claims that all nouns in classifier languages are mass nouns. However, these arguments were also met with some challenges. For example, Cheng and Sybesma (1999) and Li (1999) provide evidence for the distinction between count and mass nouns in Chinese based on their different syntactic features, while Imai and Mazuka (2007) and Her and Chen (2013) propose in their empirical studies that the ontological or lexical distinction between count and mass nouns (or objects and substances) is universally present in all languages no matter whether it is marked syntactically or not. However, the role of numeral classifiers in numeral noun phrases is not in dispute. According to Cheng and Sybesma (1999: 515), numeral classifiers create for their noun referents a unit either in terms of measure or 'natural semantic partitioning'. As indicated in Example (23), the nouns *jiǔ* 'liquor' and *bǐ* 'pen' in Chinese are individuated by *píng* 'bottle' as a unit of measure in (a) and by *zhī* 'CLF:STICK-LIKE' as a unit of 'natural semantic partition' or a semantic unit in (b).

(23) Numeral classifiers in Chinese (adapted from Cheng and Sybesma 1999: 514)
    a. *sān*   *píng*   *jiǔ*
       three   bottle   liquor
       'three bottles of liquor'
    b. *sān*   *zhī*   *bǐ*
       three   CLF:STICK-LIKE   pen
       'three pens'

Thirdly, classifiers can also be used to express a speaker's attitudes toward the referent. As classifiers have meanings related to humanness and animacy, and also kinship or social status, change of classifiers can convey affection or contempt. For example, in Jacaltec (Mayan), the noun classifiers *naj* 'CLF:MALE NON-KIN' and *ix* 'CLF:FEMALE NON-KIN' can convey an insult with reference to an old or famous person who is expected to be referred to by the noun classifier *ya7* 'CLF:RESPECTED HUMAN'. On the other hand, if *ya7* is used to refer to someone normally classified as *naj* 'CLF:MALE NON-KIN' or *ix* 'CLF:FE-MALE NON-KIN', it can show a compliment (Craig 1986a: 270). In Burmese, numeral classifiers can ascribe specific properties to nouns denoting animacy. For example, saints are categorized by *'pa* 'CLF:CLOSE', people of higher status are assigned to the classifier *'u*

'head', ordinary humans are classified by *yauʔ* 'CLF:PERSON', while nouns denoting animals occur with *kauŋ* 'CLF:BODY' (Becker 1975: 116). The social status of the referent can also be upgraded or degraded by change of numeral classifiers. For example, animal referents can be upgraded by being reassigned with *'pa* 'CLF:CLOSE' and human referents can be downgraded by being reassigned with *kauŋ* 'CLF:BODY' (Becker 1975: 115).

Like gender, all types of classifiers can play an important part in the organization of discourse. The presence of classifiers and the ordering of elements in a classifier phrase can be used to establish the status of nominal referents, while the choice of different classifiers can be used to manipulate the status of the referent by presenting it from different perspectives. The establishment and manipulation of the status of a referent is usually referred to as 'referent tracking', which is regarded as the primary function of classifiers in discourse (Aikhenvald 2000: 329-333). To be more specific, classifiers have three discourse functions: reference identification, reference management, and re-presentation of referents (Contini-Morava and Kilarski 2013: 279-291).

First, classifiers can be used anaphorically to identify noun referents, even though they do not occur in concordial agreement, as gender markers do. As indicated in example (24) in Jacaltec (Mayan), noun classifiers *naj* 'CLF:MALE' and *no7* 'CLF:ANIMAL' in (b) are used to identify their referents, which are referred to explicitly in (a). According to Craig (1986a: 264), such classifiers can be used as anaphoric pronouns in Jacaltec.

(24) Noun classifiers in Jacaltec (Craig 1986a: 264)
   a. *xil*   *naj*   *xuwan*   *no7*   *lab'a*
      saw   CLF:MAN   John   CLF:ANIMAL   snake
      'John saw the snake'
   b. *xil*   *naj*   *no7*
      saw   CLF:MALE   CLF:ANIMAL
      'He saw it'

Second, classifiers are used to indicate definiteness, referentiality and topicality (Contini-Morava and Kilarski 2013: 283-284). This is one of the typical functions of classifiers, as they are less grammaticalized and thus more optional in terms of their presence, choice, and ordering in classifier phrases. For example, numeral classifiers, such as in Cantonese (Sino-Tibetan), can be used to express definiteness and specificity (Cheng and Sybesma 2012b). As indicated in example (25), the numeral classifier *zek* in (b) expresses definiteness and corresponds to the article *the* in English. The two examples in (25) show that the

placement of classifier phrases can affect the interpretations of numeral classifiers as definite or indefinite. Compared with the numeral classifier in a pre-verbal classifier phrase in (b), *bun* 'CLF:VOLUME' in a post-verbal classifier phrase in (a) expresses indefiniteness.

(25) Numeral classifiers in Cantonese (Sino-Tibetan) (Cheng and Sybesma 2012b: 16)
   a. *ngo*   *soeng*   *maai*   *bun*   *syu*   *lei*   *taai.*
       I   want   buy   CLF:VOLUME   book   come   read
       'I want to buy a book (to read).'
   b. *zek*   *gau*   *soeng*   *gwo*   *maalou.*
       CLF   dog   want   cross   road
       'The dog wants to cross the road.'

Finally, classifiers can also help recategorize noun referents in discourse. The re-presentation of referents involves a change in the use of classifiers and thus a change in perspective on the referents. Example (26) comes from a narrative text in Toba (Guaicuruan; Argentina). The narrator changes the perspective on the referent of the noun *wallikyaGay* 'capybara' by using two different deictic classifiers: *na* 'CLF:PROXIMAL' and *so* 'CLF:DISTAL'. The classifier *na* 'CLF:PROXIMAL' in (a) is used to introduce and direct attention to the referent of *ana wallikyaGay* 'the capybara', while *so* 'CLF:DISTAL' in (b) is used to represent the known referent and establish it as the topic of the story.

(26) Deictic classifiers in Toba (Guaicuruan) (Messineo and Cúneo 2019: 210)
   a. *nache*   *so*   *y-alliʔi*   *s-oGonaGana-q*
       CONJ   CLF   POSS1-brother-in-law   1A-hunt-PL
       'Then [one day] [with] my brother-in-law we hunted'
       ñ-aqataGañi   *a-na*   *wallikyaGay*   *na* = Focus
       1MAL-catch   FEM-CLF   capybara
       'we caught a CAPYBARA'[5]
   b. *nache*   *ayem*   *i-nak*   *a-so*   *wallikyaGay*   *so* = Topic
       CONJ   PRON1   3A-bite   FEM-CLF   capybara
       'then the capybara bit me' (lit: 'then it bit me, the capybara')
       ayem   *y-acha-ngi*   *ʔetaxat*
       PRON1   3A-catch-DIR   water
       'pulled me into the water'

In summary, classifiers can serve two main functions: semantic, where they can be used to individuate noun referents, specify and ascribe properties, and discourse, where they

---

[5] Capital letters indicate focused information.

are used to establish and manage the status of referents. These functions can also be found in gender systems, except that the semantic function involving lexical expansion is more typical in gender systems with overt morphological marking on nouns, while the discourse function concerning reference management is more typical of classifiers, which can be used to express definiteness, specificity and topicality in discourse due to their optional nature.

2.5. Diachrony of nominal classification

This section will give a brief introduction to the diachrony of nominal classification. I will discuss the changes in gender and classifiers, typically concerning their origin, development, and loss.

Gender can develop either 'from above' through grammaticalization of less obligatory forms, e.g., classifiers, or 'from below' through reanalysis of existing morphosyntactic patterns, e.g., pronouns and derivational markers (Luraghi 2011). The grammaticalization of noun classifiers to gender can be shown in the three steps of a grammaticalization chain: 1) a generic-specific 'pairing' (a generic noun classifier and a specific noun), 2) the repetition and sometimes omission of the obligatory use of noun classifiers, and 3) morphological and phonological erosion of noun classifiers (Aikhenvald 2000: 372-373; 2016a: 79-81). For example, the noun classifiers *nà-* 'feminine' and *dà-* 'masculine' in Mupun (West Chadic, Afro-Asiatic; Nigeria) are derived from the nouns *nàa* 'woman' and *dàa* 'man' used in vocative expressions. They appear on proper names, some common nouns, as well as a pronominal element *kómtàk* 'such and such', as in *nà-kómtàk* 'such and such a female' and *dà-kómtàk* 'such and such a male' (Frajzyngier 1993: 49-53; Aikhenvald 2016a: 79). The noun classifiers *é-* 'CLF:ANIM' and *a=* 'CLF:ANIM' in Ngan'gityemerri (Southern Daly; Australia), on the other hand, have developed some features of agreement, as they occur on nouns and their modifiers, as shown in *é-melpe a=yéyi* (CLF:ANIM-stingray CLF:ANIM=other) 'another stingray') (Reid 1997: 215-217).

Gender can also develop from existing morphosyntactic patterns such as derivational markers, pronouns as well as number and case marking. The development of gender from pronouns has been demonstrated in several studies. Greenberg (1978) proposed that

the development of gender markers undergoes the following stages: a demonstrative (or a 3rd person pronoun) > definite article > non-generic article > noun marker. For example, the gender marker *kɔ́* 'he (masculine)' in Zande (Ubangi, Niger-Congo; DR of Congo) developed from the 3rd personal pronoun *kɔ́* 'he (masculine)' (Aikhenvald 2016a: 76-77). As illustrated in example (27), *kɔ́* is shown in agreement as a prefix on the verb *ni* 'be' and a suffix attached to the adjective *bakɛ́rɛ́* 'big'.

(27) Gender marker *kɔ́* in Zande (Ubangi, Niger-Congo) (Aikhenvald 2016a: 77)
    kɔ́-ni   *bak*ɛ́rɛ́*-kɔ́*
    he-be   big-he
    'He is big'

Gender derived from number marking may result in a human-based distinction in a genderless language. For example, the gender marker *kadag* indicating group plurality occurs only with humans in Balochi (Western Iranian, Indo-European; Pakistan) (Aikhenvald 2016a: 82). Gender can also originate from case marking, as illustrated by the development of gender in Proto-Indo-European (Luraghi 2011).

    Classifiers are most likely to develop through grammaticalization of nouns. Aikhenvald (2000: 354) distinguishes five groups of nouns as the sources for different types of classifiers: nouns for body parts, nouns referring to kinship, humans and higher animates, generic nouns, unit counters, and culturally important items. For example, numeral classifiers in some African languages originated from body part nouns, e.g., the classifier for small globular objects from the noun 'eye' in Denya (Atlantic-Congo; Cameroon) and the classifier for inanimate objects from 'body' in Busa (Mande; Nigeria) (Kießling 2018: 39-42). Noun classifiers typically develop from phonologically reduced nouns for humans and animals, as shown in the following examples from Mam (Mayan; Guatemala and Mexico): *jal* 'nonhuman' < *jiil* 'wild animal', *q'a* 'young man' < *q'aa* 'young man', *txin* < *txiin* 'young woman', *ma* < *matiij* 'big', *xu7j* < *xu7j* 'woman' (England 1983: 158). Verbs are a rarer source for classifiers. Posture and motion verbs can develop into classificatory verbs, deictic and verbal classifiers, and verbs of handling can be grammaticalized into verbal, possessed, relational, and numeral classifiers (Aikhenvald 2000: 362). For example, deictic classifiers in Siouan languages mentioned in §2.4.1.5 derived from the verbs 'sit', 'stand', 'lie' and 'move' (Rankin 2004: 211-212), while some verbal classifiers in Imonda, a Border language of Papua New Guinea

developed from the verbs of handling *põt-* 'pick fruit', *puiual* 'break in two', and *i* 'scoop water out' (Seiler 1986: 190-193).

Grammaticalization paths in classifier systems can be complex by involving the polygrammaticalization of different types of classifiers deriving from the same open lexical classes, the coexistence of classifiers showing different degrees of grammaticalization, and semantic changes and extensions. Polygrammaticalization in Palikur (Arawakan; Brazil) can be illustrated by the development of the noun *kig* 'nose' into the verbal classifier *-kig* 'CLF:POINTED OBJECTS' and the locative classifier *-kigsa* 'CLF:ON.POINTED' (Aikhenvald 2000: 193, 357, 375). Classifiers of different degrees of grammaticalization are found in Ngan'gityemerri (Southern Daly; Australia), where noun classifiers vary in terms of morphosyntactic bondedness and have developed inflectional agreement typical of the most grammaticalized system of gender (Reid 1997: 215-217; cf. Seifart 2010: 727-728). As shown in example (28), *gagu* 'CLF:ANIMAL' in (a) is an independent element, *wa=* 'CLF:MALE' in (b) is a proclitic, and *é-* 'CLF:ANIM' and *a=* 'CLF:ANIM' in (c) are prefixes marked on the noun and its modifier as agreement markers.

(28) Noun classifiers in Ngan'gityemerri (Southern Daly; Australia) (Reid 1997: 215-217)
    a. *gagu*   *wamanggal-kɔ́*   *gagu*   *kerre*   *ngeben-da*
        CLF:ANIMAL   wallaby   CLF:ANIMAL   big   1SG.AUX-shoot
        'I shot a big wallaby'
    b. *wa=ngurmumba*   *wa=ngayi*   *darany-fipal-nyine*
        CLF:MALE=youth   CLF:MALE=mine   3SSG.AUX-return-FOC
        'My initiand son has just returned'
    c. *é-melpe*   *a=yéyi*
        CLF:ANIM-stingray   CLF:ANIM=other
        'another stingray'

Classifier systems also show semantic changes and extensions. It is generally acknowledged that the meanings of classifiers are usually extended from concrete to more abstract, as illustrated by the grammaticalization of the nouns for 'stalk/tree', 'fruit', and 'leaf' to numeral classifiers indicating the three basic shapes of long, round, and flat (Adams and Conklin 1973: 5). Recent studies also show that classifiers undergo a semantic reduction shown in the following process: open lexical nouns > specific classifiers with transparent semantics > general classifiers with more opaque semantics (Seifart 2018: 20-23). According to Zubin and Shimojo (1993: 491), general classifiers tend to be unspecified in

terms of semantics, and thus, they are used as default classifiers complementary to specific classifiers.

The development of nominal classification systems can also be attributed to an external motivation. The role of areal diffusion can be illustrated by the existence of similar types of nominal classification in different languages in geographically contiguous zones and complex systems of nominal classification in languages in areas where different language families meet. For example, gender is predominant in the languages of Africa and Indo-European languages, while numeral classifiers occur in most languages in East and Southeast Asia. On the other hand, gender and classifiers have been found to coexist in such linguistic areas as north-eastern and southern Amazonia (Aikhenvald 2012: 300-303) as well as South Asia (Allassonnière-Tang and Kilarski 2020) at the meeting point of Indo-European and Sino-Tibetan languages. For example, the cooccurrence of gender and numeral classifiers in Nepali can be attributed to the spread of numeral classifiers from Tibeto-Burman languages to Indo-Aryan languages with gender systems (Allassonnière-Tang and Kilarski 2020), while numeral classifiers in Southeast Asia appeared as a result of the language contact with classifier languages in Tai family (cf. Emeneau 1956; Barz and Diller 1985; Bisang 1996). It is also the case with numeral classifiers in East Asia, although a debate has focused on whether they are an areal language feature that spread from Tai to Chinese. A number of scholars suggested that numeral classifiers in East Asia developed due to areal diffusion from Tai (e.g., Jones 1970; Erbaugh 1986; Peyraube 1991). However, most linguists in China argued that numeral classifiers are indigenous to Chinese (e.g., Wang 1994: 168-169; Wu 2014a; Bu 2011a, 2011b). A recent study by Her and Li (2019) proposed that numeral classifiers in languages of Asia and the Pacific were acquired through language contact with most probably Sinitic languages.

Reduction and loss of nominal classification systems can be attributed to internal developments or, more frequently, to external motivation as a result of language contact. For example, the loss of grammatical gender in English took place along with the decay of case and agreement (Aikhenvald 2016b). On the other hand, some varieties of northwest Mandarin retained only the general classifier and lost most of their specific classifiers due to language contact with non-Sinitic Altaic languages, such as Mongolic, Turkic and Tungusic languages (Sandman and Di Garbo In press).

In conclusion, devices of nominal classification develop through grammaticaliza-tion from open lexical classes, e.g., nouns and verbs, or from reanalysis of more closed word classes and grammatical categories, e.g., pronouns, derivational affixes and case marking. Therefore, they vary in terms of the degree of grammaticalization, morphosyn-tactic features and semantic complexity. Furthermore, the development of nominal clas-sification systems can also be attributed to language contact, and thus, they may charac-terize large linguistic areas. Finally, the loss of gender and classifiers in some languages can be motivated internally or externally due to language contact.

2.6. Overview of the study on nominal classification

Gender is regarded as "classes of nouns reflected in the behaviour of associated words" (Hockett 1958: 231), while classifiers are defined as morphemes "in surface structure under specifiable conditions" to denote "some salient perceived or imputed characteristic of the entity to which an associated noun refers (or may refer)" (Allan 1977: 285). They are regarded as important devices to classify noun referents.

Gender in Indo-European languages has long been the focus of attention. Accord-ing to Kilarski (2013: 59-60), the accounts of gender can be traced back to c. 2500 years ago, when the terminology of gender and their core descriptions concerning form and meaning correlation, agreement and assignment were established. For example, Protago-ras (c. 480- c. 410 B.C.) was reported to make the first reference to gender (cf. Aristotle, Rhet. 3.5), while Aristotle was first to apply formal criteria to the classification of nouns (Poetics 1458a 9–16). Apollonius Dyscolus (c.110–175 A.D.) described gender in terms of agreement (cf. Blank 1982). In contrast, Ammonius Hermiae (ante 445-517/526 A.D.) gave an account of gender with regard to assignment (cf. Blank 1996). While gender in Arabic and Hebrew was already studied in the Middle Ages, more studies on these lan-guages as well as other non-Indo-European languages were carried out around the 17th century. For example, noun classes were recorded in the first grammar of a Bantu lan-guage, Kongo, spoken in present-day Angola and the Democratic Republic of the Congo (Brusciottus 1659).

Similarly, numeral classifiers in East Asia were recorded in the 17th and 18th cen-tury missionary grammars, in which numeral classifiers in Mandarin Chinese were

recognized as *particulas* (i.e. particles) by Varo (1703) (Coblin and Levi 2000), and their semantic correlation with noun referents was also identified (cf. Kilarski 2013: 108-109). Little attention was devoted to classifiers up till the mid-20th century. In the 19th century, further numeral classifier languages in East Asia and North America were identified (Alcock 1861:36). Most of these systems were described as communicatively useless or semantically redundant (Brinton 1885: 62; Greenberg 1972: 330; Lehman 1979: 174). However, these findings, typically concerning the classificatory verbs in Cherokee, a Southern Iroquoian language spoken today in Oklahoma and North Carolina, were also the basis for subsequent analyses of their lexical and structural properties by Brinton (1885), and more notably, for the analysis of their semantics by Hewitt (1893).

More systematic work on nominal classification began in the 1970s. A series of studies in the 1970s showed that classifiers are semantically and pragmatically motivated (e.g., Friedrich 1970; Adams and Conklin 1973; Adams et al. 1975; Becker 1975; Denny 1976; Allan 1977). In the earliest survey of numeral classifiers in Asian languages, Adams and Conklin (1973) discussed their semantic properties and suggested that there are semantic universals based on the visual feature of shape. Denny (1976) focused on the three types of interaction with classifiers: physical interaction, function interaction, and social interaction. Research in this period also dealt with semantic universals in classifier systems (e.g., Sanches and Slobin 1973), which contributed to comparisons of different nominal classification systems.

The 1970s and 1980s are characterized by early work on the typologies of nominal classification. For example, based on morphosyntactic properties, Allan (1977) proposed four types of classifiers: numeral classifiers, concordial classifiers (noun classes), predicate classifiers (verbal classifiers), and intra-locative classifiers (deictic classifiers). Dixon (1982c) contrasted the grammatical category of noun classes and the semi-open lexical-like systems of classifiers and illustrated the development of classifiers into noun classes. Serzisko (1982: 95) arranged gender, noun class, and numeral categorization on a "scale of classificatory techniques", and compared them based on parameters such as grammaticality, semantic complexity, and variability.

The typologies of nominal classification that have been proposed were motivated by new data that became available as well as considerable ambiguity in terminology. Specific studies include Aikhenvald (2000, 2004, 2012, 2017) and Grinevald (2000, 2001, 2002, 2003, 2004). Aikhenvald's *Classifiers: A typology of noun categorization devices*

(2000) remains the most comprehensive work on systems of nominal classification. In this book, Aikhenvald (2000) distinguished among them in terms of such criteria as morphosyntactic loci and degree of grammaticalization and also discussed the functions of nominal classification. Grinevald's morphosyntactic typologies of classifiers were established based on the continuum of grammaticalization (Grinevald 2000). She also suggested a prototype approach to classifier systems, i.e., types of nominal classification are distinguished based on "the most contrastive characteristics" (Grinevald 2000: 80). Based on these studies, Contini-Morava and Kilarski (2013) proposed a functional typology of nominal classification, and argued that nominal classification devices have semantic and discourse functions.

Recent approaches to the typologies of nominal classification focus on the canonical properties of individual types of nominal classification, in particular gender. Corbett and Fedden established the extremes of nominal classification systems, from the simplest to the most complex ones, based on semantic and formal assignment and related morphological means (Corbett 2006; Corbett and Fedden 2016; Fedden and Corbett 2017; Fedden et al. 2018). These studies deal with cooccurring systems of classifiers and genders in one language, and cross-linguistic comparisons of the same type of nominal classification. The typologies of nominal classification proposed in this period are also closely related to the notion of language complexity. Corbett (2006) proposed three principles of canonical agreement of gender based on information content (redundant vs. informative), syntax (simple vs. complex), and morphological realization (inflectional vs. lexical). Audring (2017) put forward two-step approaches to calibrate the complexity of grammatical gender based on the different dimensions of gender and their different degrees of complexity. The notion of covert complexity related to multifunctionality proposed by Bisang (2014) contributes to the understanding of the role of classifiers in such languages as Chinese in measurement of language complexity.

Other than the typologies of nominal classification, nominal classification systems have also been approached from other perspectives. These include language acquisition of, e.g., Chinese numeral classifiers (e.g., Erbaugh 1986; Hu 1993; Erbaugh 2006; Li et al. 2008; Huang and Chen 2014). Gender and classifiers have also been approached in psycholinguistic studies with the aim to find out how they are processed and retrieved by speakers of such languages as Chinese, Japanese, and Korean, in different cultures and social backgrounds (e.g., Saalbach and Imai 2012; Xu et al. 2013; Kemmerer 2017; Jin

2018; Her et al. 2018). Furthermore, diachronic accounts of the development of nominal classification have also been provided, including the role of language contact (e.g., McGregor and Wichmann 2018; Passer 2016). Allassonnière-Tang and Kilarski (2020) described the co-occurrence of gender and numeral classifiers in Nepali (Indo-European, Indic), while Sandman and Di Garbo (In press) found a dramatic reduction of numeral classifiers and the co-existence of gender in several varieties of northwest Mandarin.

In summary, philosophers and linguists have long devoted attention to the phenomenon of nominal classification. While previous studies have examined their properties as well as other aspects related to language acquisition and processing, there is need for further studies, e.g., on cross-linguistic comparisons concerning form and meaning correlation as well as the functions of specific devices of nominal classification in specific languages.

2.7. Concluding remarks

This chapter has outlined the state of research on nominal classification systems, focusing on their semantics, functions and expression. Nominal classification systems can be divided into gender and classifiers based on the degree of grammaticalization, assignment principles and morphosyntactic expression. Gender can be distinguished from classifiers by its definitional property of agreement, while classifiers can be further divided into subtypes based on different morphosyntactic expression. Studies on nominal classification systems have approached them from different perspectives such as acquisition, linguistic complexity and language processing. However, relatively little has been done on the comparison of languages with different types of nominal classification, e.g., a classifier language in opposition to a non-classifier language, typically based on their functionality. Furthermore, more studies should also be carried out based on large corpora with rich discourse data. In the following chapters, I will first review the features and functions of Chinese numeral classifiers, before conducting a corpus-based comparison between numeral classifiers in Chinese and their corresponding forms in English based on their functionality.

Chapter 3: Chinese numeral classifiers

3.1. Introduction

Chinese has a rich system of numeral classifiers. They appear as independent morphemes and characters of basic individual graphic units between numerals or demonstratives and nouns. They tend to denote humanness, animacy and physical properties of noun referents, and are used to individuate nouns, specify the properties of their referents, and track referents in discourse. This chapter deals with Chinese numeral classifiers in terms of their features, functions as well as their diachronic development.

In this chapter, I will first give an overview of the syntactic and semantic features of Chinese numeral classifiers, as well as the distinctions between numeral classifiers and measure words in §3.2. In §3.3, I will examine the functions of Chinese numeral classifiers, which will be the foundation of the discussions on the comparison of Chinese numeral classifiers with their English equivalent forms in the translation between the two languages in the following chapters. §3.4 gives an account of the diachrony of Chinese numeral classifiers. Before the conclusion in §3.6, a brief literature review will be made of different approaches to Chinese numeral classifiers in Chinese and Western linguistics in §3.5.

3.2. Syntactic and semantic features of Chinese numeral classifiers

The obligatoriness of Chinese numeral classifiers and the structure of numeral noun phrases will be examined first, and then the semantic features of different types of Chinese

numeral classifiers will be discussed. Based on their syntactic and semantic features, a distinction between numeral classifiers and measure words will be made in §3.2.3.

3.2.1. Syntactic features

Chinese numeral classifiers occur as obligatory elements in numeral noun phrases. Numeral noun phrases in Chinese are usually composed of three basic elements: numerals, demonstratives or interrogative pronouns, numeral classifiers/measure words, and nouns. As shown in example (29), Chinese numeral classifiers occur next to numerals, e.g., *yī* 'one', *liǎng* 'two', *sān* 'three' in (a), (c) and (d) respectively, or such demonstratives as *zhè* 'this' in (b), or such interrogative pronouns as *ná* 'which' in (e).

(29) Chinese numeral noun phrases

    a. *yī*   *běn*   *shū*
        one   CLF:BOOK   book
        'one book'

    b. *zhè*   *xiāng*   *shū*
        this   MENS:BOX   book
        'this book'

    c. *liǎng*   *dà*   *gè*   *píngguǒ*
        two   big   CLF:GENERAL   apple
        'two big apples'

    d. *sān*   *dà*   *lán*   *píngguǒ*
        three   big   MENS:BASKET apple
        'three big baskets of apples'

    e. *ná*   *gè/lán*   *píngguǒ*
        which   CLF:GENERAL/MENS:BASKET   apple
        'which basket of apples'

Numeral classifiers and measure words take the same slot in numeral noun phrases and can be both preceded by adjectives. As shown in example (29), the adjective *dà* 'big' can be found as a modifier of the numeral classifier *gè* 'CLF:GENERAL' in (c) and the measure word *lán* 'MENS:BASKET' in (d). However, they are translated into modifiers of different elements in English in the two examples: *dà* 'big' in example (c) is translated into a modifier of the noun *apples*, while it is a modifier of the measure word *basket* in (d). This will be further analysed in §3.2.3. Nouns usually occur as the head in numeral noun phrases and can be preceded with modifiers. The dissertation deals with semantics and functions

of Chinese numeral classifiers as shown by the interplay of numerals, typically *yī* 'one', numeral classifiers, nouns, and modifiers of numeral classifiers and nouns. Therefore, demonstratives and interrogative pronouns will not be examined in this study.

The obligatoriness of Chinese numeral classifiers is usually ascribed to the unspecified nature of nouns in Chinese. Chinese nouns have been regarded as mass nouns (e.g., Chierchia 1998b; Jin 2013), and can only be counted based on their semantic unit. However, this view was challenged by Cheng and Sybesma (1999: 515), who argued that there is a count/mass distinction in Chinese nouns and the count nouns among them are analogous with English count nouns without number morphology (Cheng and Sybesma 1999: 519-520). They argue that count and mass nouns in Chinese can be distinguished by different ways of enumeration: count nouns are enumerated by count classifiers (numeral classifiers) by naming the "unit of natural semantic partitioning", while mass nouns are enumerated by mass-classifiers or massifiers (measure words) in terms of "unit of measure" (Cheng and Sybesma 1999: 515). As illustrated in example (30), *shū* 'book' in example (a) can be discretely individuated, so it can be counted in terms of semantic unit created by *běn* 'CLF:BOOK', while in example (b), *shuǐ* 'water', as a mass noun that cannot be separated in terms of semantics, is enumerated by the measure word *bēi* 'MENS:CUP' as a unit of quantity. Zhang (2012) further points out that non-mass nouns in Chinese are counted by numeral classifiers as a unit of quality in terms of delimitive features, e.g., shape, size and dimensions (cf. Zhang 2012: 11-13). As indicated in example (c), the noun *málù* 'road' is enumerated by the numeral classifier *tiáo* 'CLF:SLENDER' based on the shape of the referent. Therefore, numeral classifiers or measure words are obligatory in Chinese numeral noun phrases to create a semantic unit either in terms of quality or in terms of quantity to enable the quantification of nouns.

(30) Numeral nouns and measure words as obligatory elements in Chinese numeral noun phrases

a.	*yī*	*běn*	*shū*
	one	CLF:BOOK	book
	'one book'		
b.	*yī*	*bēi*	*shuǐ*
	one	MENS:CUP	water
	'one cup of water'		
c.	*yī*	*tiáo*	*málù*
	one	CLF:SLENDER	road
	'one road'		

Chinese numeral noun phrases have been regarded as either right-branching, or a left-branching, or a mixture of both. According to Tang (1990), Cheng and Sybesma (1998), and Li (1999), among others, Chinese numeral noun phrases present a unified right-branching structure. In their interpretation, numeral classifiers, e.g., *běn* 'CLF:BOOK', and measure words, e.g., *bēi* 'MENS:CUP', form a constituent with nouns, as indicated in Fig. 7.

NUM	CL/M	N
yī	*běn*	*shū*
one	CLF:BOOK	book
'one book'		
yī	*beī*	*shuǐ*
one	MENS:BOX	water
'one box of books'		
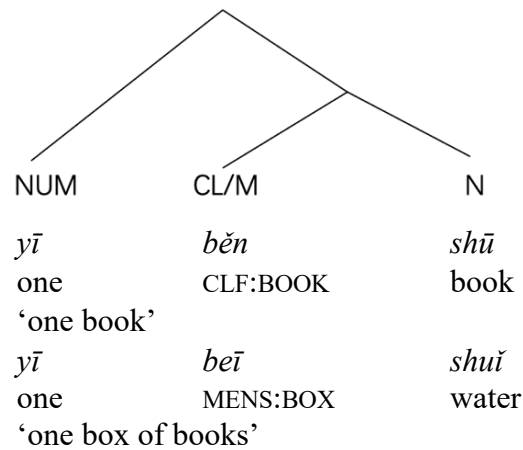
Fig. 7. Right-branching structure of Chinese numeral noun phrase.

In another interpretation, Chinese numeral noun phrases are regarded as left-branching. According to Greenberg (1972: 185), numeral classifiers, like measure words, form a unit with numerals and form the constituent of [Num+CLF]. More recent studies have corroborated this proposal by claiming a unified left-branching structure in Chinese, e.g. Hsieh (2008), Her (2017), and Tang et al. (2021). In their view, numerals and numeral classifiers or measure words form a constituent before it modifies head nouns, as indicated in Fig. 8.

NUM	CL/M	N
yī	*běn/xiāng*	*shū*
one	CLF:BOOK/	book
	MENS:BOX	
'one book/one box of books'		
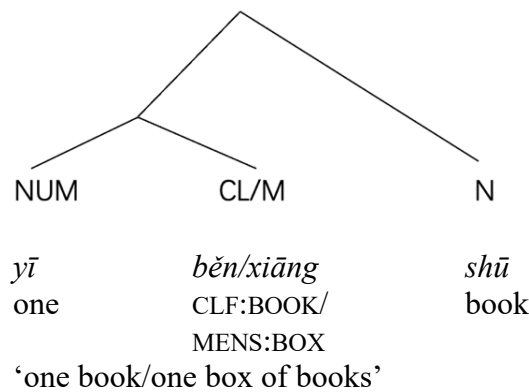
Fig. 8. Left-branching structure of Chinese numeral noun phrase.

Another interpretation combines both views, or more specifically, numeral noun phrases with measure words are right-branching, while numeral noun phrases with numeral classifiers are left-branching (e.g., Zhang 2012: 143-147; Li 2013: 153).

The dissertation follows Her (2017) and Tang et al. (2021), and argues that Chinese numeral noun phrases have a left-branching structure. Her (2017) proposed that numeral noun phrases are left-branching structures from both diachronic and synchronic perspectives. His diachronic studies demonstrate that the constituent [NUM+CLF] has been found in Chinese history for about 3000 years, while his synchronic study shows that in some typologically different languages in the Sino-Tibetan family and 52 genetically different classifier languages, numeral classifiers form a constituent with numerals before they are used to modify nouns (Her 2017). Tang et al. (2021) also find evidence at a statistically significant level in two psycholinguistic experiments that numeral noun phrases in Chinese are left-branching. Aside from evidence provided by Her (2017) and Tang et al. (2021), evidence concerning the origin of Chinese numeral classifiers also shows that numeral noun phrases in Chinese are left-branching. According to Bu (2011b) and Feng (2012: 91), the origin of Chinese numeral classifiers was triggered with the disyllablization of Classic Chinese and they are used with numerals to form a minimal independent prosodic unit with at least two syllables.

In conclusion, numeral classifiers are obligatory elements occurring between numerals or demonstratives and nouns. They create a unit of quantity to enable the quantification of nouns and thus numeral noun phrases are left-branching.

3.2.2. Semantic features

While Chinese numeral classifiers are obligatory in numeral noun phrases, their choice is semantically motivated. In this section, I will focus on the semantics of Chinese numeral classifiers, including general and specific classifiers and their semantic parameters.

The system of Chinese numeral classifiers consists of several general classifiers and a large inventory of specific classifiers. General classifiers can cooccur with a wide range of nouns, while specific classifiers can only be used for a limited group of referents. Furthermore, general classifiers differ from specific classifiers in discourse. While specific classifiers tend to be used to introduce a new referent, general classifiers can replace

them in subsequent mentions. This point will be described in more detail in §3.3.2. The most frequently used general classifier in Modern Chinese is *gè* 'CLF:GENERAL', and other general classifiers include *méi* 'CLF:GENERAL, ROUND.PIECE', used as a general classifier in ancient Chinese between the first and the sixth century (Wu 2014b: 64-71), and *zhǒng* 'CLF:KIND, GENERAL' described as a general classifier by Huang and Ahrens (2003: 17-20). I will come back to the development of general classifiers in §3.4.1.

As mentioned above, *gè* is the most frequently used general classifier in Modern Mandarin Chinese. It is often the first numeral classifier acquired by children (e.g., Hu 1993: 59-62; Erbaugh 2006), and it is the default classifier among non-native speakers (e.g., Zhang and Lu 2013). It has also been reported to be the only numeral classifier retained in language contact situations involving attrition, as in Wutun (Sandman and Di Garbo in press). However, the classifier cannot be used with all nouns, as illustrated in recent studies that have shown restrictions in its use. For example, Zhou (2014: 91) and Frankowsky and Ke (2016: 63-65) point out that it is less likely to be used with nouns denoting plants, e.g. *shù* 'tree', *huā* 'flower' and *cǎo* 'grass', animals depending on their perceived distance from humans, e.g. *gǒu* 'dog' and *niǎo* 'bird', or animals with such salient physical properties as length, e.g. *niú* 'ox, cow' and *shé* 'snake'.

The numeral classifier *zhǒng* is the second most frequently used numeral classifier.[6] It is used to individuate nouns in terms of kind or type instead of unit (Huang and Ahrens 2003). Its use is illustrated in example (31). While *xuéshēng* 'student' refers to an individual student when used with *gè*, it refers to a group of students of the same kind when used with *zhǒng* as indicated in (a). Furthermore, *zhǒng* is more likely to be the default numeral classifier for those nouns referring to a certain kind of product or material, e.g., *kàngshēngsù* 'antibiotic' as in (b).

(31) The use of *zhǒng* to individuate referents in terms of kind
    a. *yī*   *zhǒng*   *xuéshēng*
          one   CLF:KIND, GENERAL   student
          'one kind of students'
    b. *yī*   *zhǒng*   *kàngshēngsù*
          one   CLF:KIND, GENERAL   antibiotic
          'one antibiotic'

---

[6] However, *zhǒng* is not treated as a general classifier, e.g., by Erbaugh (1986), Wu (2014b), and Gao and Malt (2009).

Subordinate to *zhǒng,* there are several other kind classifiers, including *yàng* 'CLF:KIND, SHAPE AND APPEARANCE' and *kuǎn* 'CLF:KIND, DESIGN'. These classifiers can only be replaced by *zhǒng* instead of *gè.*

In contrast to general classifiers, specific classifiers are used to denote specific properties of noun referents. Similar to other systems of numeral classifiers discussed in §2.4.1.1, Chinese numeral classifiers categorize noun referents in terms of humanness, animacy, or physical properties, typically shape.[7]

Chinese numeral classifiers categorize noun referents in terms of humanness. While the general classifier *gè* is regarded as the default classifier for nouns denoting humans, there are also other specific classifiers for humans. For example, *wèi* 'CLF:INDI-VIDUAL, RESPECT' can be used with reference to people who are of a higher social status. The use of *gè* and *wèi* for humans is illustrated in example (32). While *gè* can only be used with the general noun *rén* 'person', as in (a), both *gè* and *wèi* can be used with *guānyuán* 'official', as in (b) and (c). However, *wèi* and other specific classifiers for humans cannot be used with the noun *rén* 'person', unless a modifier is present.

(32) Chinese numeral classifiers for humans
  a.   *yī*   *gè*   *rén*
       one   CLF:GENERAL   person
       'one person'
  b.   *yī*   *gè*   *guānyuán*
       one   CLF:GENERAL   official
       'one official'
  c.   *yī*   *wèi*   *guānyuán*
       one   CLF:INDIVIDUAL, RESPECT   official
       'one official'

While some numeral classifiers are used exclusively with nouns for humans, other classifiers are used for animals and inanimates. The most typical one is *zhī* 'CLF:SINGLE', which can be regarded as a general classifier for animals (Tai 1992: 594).

Other numeral classifiers for animates include *tóu* 'CLF:HEAD', *pǐ* 'CLF:SIN-GLE.HORSE', and *tiáo* 'CLF:SLENDER', which are usually used to specify a particular feature of their referents. As shown in example (33), *tóu* 'CLF:HEAD' is used with animals with a head with salient features, e.g., large size, as with *nǎiniú* 'cow' in (a), *pǐ* with nouns

---

[7] For recent descriptions of the semantics of Chinese numeral classifiers, see, e.g., Tai (1992: 594-601), Gao and Malt (2009: 1171-1177), Wu (2014b: 63-136), and Song (2017: 43-161).

referring to horses or horse-like animals as in (b), and *tiáo* for animals with a slender shape, e.g., *gǒu* 'dog' in (c).

(33) Specific classifiers used for animals
    a.   *yī*   *tóu*   *nǎiniú*
         one   CLF:HEAD   cow
         'one cow'
    b.   *yī*   *pǐ*   *mǎ*
         one   'CLF:SINGLE.HORSE'   horse
         'one horse'
    c.   *yī*   *tiáo*   *gǒu*
         one   'CLF:SLENDER'   dog
         'one dog'

Other numeral classifiers are used for inanimate objects. As mentioned above, both *gè* 'CLF:GENERAL' and *zhī* 'CLF:SINGLE' are general classifiers for inanimate nouns.

However, inanimate nouns are more likely to be used with specific classifiers involving physical properties such as shape, size and other salient features. For example, *tiáo* 'CLF:SLENDER' and *kuài* 'CLF:LUMP-SHAPE' denote two-dimensional and three-dimensional shapes, respectively, while *kē* 'CLF:ROUNDISH' and *lì* 'CLF:GRAIN-LIKE' are used for small and three-dimensional objects. However, as shown in example (34), *tiáo* 'CLF:SLENDER' can also be used a wider range of nouns, including such animate nouns as *hǎohàn* 'true man' and *gǒu* 'dog' in (a) and (b) and such abstract nouns as *yìjiàn* 'opinion' in (c). In (a) and (b), *tiáo* 'CLF:SLENDER' denotes the physical shape of the two referents, while in (c), the property of being slender is metaphorically mapped to the referent based on the fact that Chinese was traditionally written vertically on a page (cf. Tai and Wang 1990: 42).

(34) Use of the numeral classifier *tiáo* 'CLF:SLENDER'
    a.   *yī*   *tiáo*   *hǎohàn*
         one   CLF:SLENDER   true.man
         'one true man'
    b.   *yī*   *tiáo*   *gǒu*
         one   CLF:SLENDER   dog
         'one dog'
    c.   *yī*   *tiáo*   *yìjiàn*
         one   CLF:SLENDER   opinion
         'one opinion'

Specific classifiers, e.g., *bǎ* 'CLF:HANDLE' and *liàng* 'CLF:VEHICLE', can also denote features based on a conspicuous part of the referents. The classifier *bǎ* 'CLF:HANDLE' collocates with nouns for objects with a handle, such as *dāo* 'knife', *sǎn* 'umbrella' and *yǐzi* 'chair'. In contrast with specific classifiers denoting shape, these classifiers are less likely to be replaced with other specific classifiers or the general classifier *gè*. Furthermore, they are also more resistant to semantic change. For example, *liàng* 'CLF:VEHICLE' has been used with *chē* 'carriage, car' since before the pre-Qin period (770-221 BC) (Ma 2015: 47-48).

Chinese numeral classifiers can also refer to an event. According to Huang and Ahrens (2003: 25-27), 'event' classifiers are used with nouns that can cooccur with such temporal delimiters as *yǐhòu* 'after', or nouns that can be collocated with such verbs as *jìnxíng* 'to proceed', *fāshēng* 'to happen', or *huā* 'to cost'. Huang and Ahrens (2003: 41) listed 35 event classifiers, including *bān* 'CLF:SHIFT' and *cháng* 'CLF:VENUE'. As shown in example (35), *bān* 'CLF:SHIFT' is used in (a) with the noun *fēijī* 'plane', where it refers to a flight rather than a plane, while in (b) *cháng* 'CLF:VENUE' refers to a scheduled screening of a movie instead of an artistic work.

(35) Chinese event classifiers
    a. *yī*   *bān*   *fēijī*
       one   CLF:SHIFT   plane
       'one flight'
    b. *yī*   *cháng*   *diànyǐng*
       one   CLF:VENUE   movie
       'one movie'

While concrete nouns in Chinese are more likely to be categorized based on physical features and thus be used with a variety of specific classifiers, the choice of numeral classifiers for abstract nouns is very limited. Abstract nouns tend to collocate only with the general classifier *gè* or the general kind classifier *zhǒng*, as in *yī gè fāngfǎ* (one CLF:GENERAL method) 'one method', *yī zhǒng tàidù* (one CLF:KIND, GENERAL attitude) 'one attitude'. Such limitations may be explained by the nature of abstract nouns, which are less easily delimitable. According to Zhou (2014: 91), the more delimitable nouns are, the more likely they cooccur with specific classifiers, and vice versa.

Chinese numeral classifiers classify their referents based on features of noun referents rather than nouns as pure linguistic forms, as discussed in §2.2. According to Lucy

(2000: 331), devices of nominal classification only classify nouns as linguistic forms when they are 'firmly established' as pure linguistic forms in gender. In Chinese, the choice of numeral classifiers is semantically correlated with their head nouns, as discussed above. While specific classifiers are referential by highlighting specific features of their referents, general classifiers are also semantically motivated, and therefore, cannot be regarded as pure linguistic forms. Furthermore, both general and specific classifiers can be used to differentiate referents and ascribe properties to their referents based on their semantics, which will be discussed in Chapter 6. However, Lucy (2000: 334-335) also acknowledged the difficulty in localizing the semantic interpretation of referents in nouns, classifiers, or the noun phrases as a whole. This issue should be addressed by examining the interplay of numeral classifiers and relevant elements in numeral noun phrases in natural language. I will come back to this point in Chapter 8.

This section has discussed the semantics of Chinese numeral classifiers. Modern Chinese has two general classifiers and a variety of specific classifiers. I summarize their semantics in Table 6 below:

Table 6. Semantics of Chinese numeral classifiers

Types	**Classifiers**	**Semantics**
entity	*gè*	general
	wèi, míng, yuán, etc.	human
	zhī	nonhumanness
	tóu, pǐ, etc.	animal
	tiáo, piàn, kuài, kē, lì, etc.	shape and size, e.g., two-dimensional, three-dimensional
	bǎ, liàng, etc.	other physical features
kind	*zhǒng*	kind, general
	yàng, kuǎn, etc.	kind, typically in shape
event	*cháng*	event, venue
	cì, bān, dùn, etc.	event, typically the duration of time

As shown in Table 6, there are three types of Chinese numeral classifiers: entity, kind and event. Entity and kind classifiers denote the same features of humanness, animacy and physical properties, as most other classifiers as discussed in §2.4.1.1, while event classifiers specify referents typically with regard to the duration of time. These semantic features are closely related to the distinction between numeral classifiers and measure words examined in §3.2.3 and are the foundation for the discussion of their functions in §3.3.

3.2.3. Chinese numeral classifiers vs. measure words

The distinction between numeral classifiers and measure words or quantifiers is essential for collecting data to be analysed in Chapter 4. This section will review previous studies to propose semantic and formal parameters to differentiate numeral classifiers from measure words in Chinese.

Both numeral classifiers and measure words occur in the same slot in numeral noun phrases in Chinese, as discussed in §3.2.1. They tended to be differentiated based on semantic parameters in previous studies. For example, according to Tai and Wang (1990: 37-38), Chinese numeral classifiers are used only with a certain number of nouns and denote inherent perceptual quantities of their referents. In contrast, measure words cooccur with a wider range of nouns and denote the quantity of noun referents. Wang (1994: 27-36) proposed six more criteria to differentiate numeral classifiers from measure words: a) *ge*-substitution: whether they can be replaced by the general classifier *gè*; b) the number of noun referents: whether they can be used to modify more than one noun referent in the same noun phrase; c) cooccurrence with *duō* 'more': whether they can be followed by the element *duō* 'more'; d) *de*-insertion: whether the genitive particle *de* can be inserted between them and nouns; e) omission in listing items: whether they can be deleted in listing items; f) use of adjectives as pre-modifiers: whether they can be pre-modified by adjectives. Among the six criteria, the first four concern semantic features. According to Wang (1994: 27-36), numeral classifiers can be replaced by *gè*, as they are all used to create a semantic unit of quality in classifier phrases. In contrast, measure words can be used to modify more than one noun referent in a noun phrase, and they can be followed by *duō* 'more', as they denote quantities of noun referents. As indicated in example (36), *kē* 'CLF:ROUNDISH' in (a) is a numeral classifier, while *xiāng* 'MENS:BOX' in (b) is a measure word referring to a container. They can be distinguished by being tested with a substitution of *gè* as in (c). The noun phrases in (a) and (c) mean the same, while the noun phrase in (b) has a different meaning in terms of the quantity of apples.

(36) The distinction between numeral classifiers and measure words
   a. *yī*   *kē*   *píngguǒ*
      one   CLF:ROUNDISH   apple
      'one apple'

b. *yī*   *xiāng*   *píngguǒ*
    one   MENS:BOX   apple
    'one box of apples'

c. *yī*   *gè*   *píngguǒ*
    one   CLF:GENERAL   apple
    'one apple'

Similarly, most measure words can also be differentiated from numeral classifiers by means of *gè*-substitution, including measure words denoting groups, e.g., *qún* 'MENS:CROWD' and *huó* 'MENS:GANG', parts, e.g., *bùfèn* 'part', arrangements, e.g., *duī* 'MENS:PILE' and *shù* 'MENS:BUNCH', weights, e.g., *dūn* 'MENS:TON', distances, e.g., *gōnglǐ* 'MENS:KILOMETRE', and periods of time, e.g., *tiān* 'MENS:DAY'.

Numeral classifiers and measure words can also be differentiated by being tested based on the number of noun referents or the cooccurrence with *duō* 'more', typically when the numeral classifiers concerned cannot be replaced by the general classifier *gè*. These numeral classifiers include specific classifiers denoting shape, e.g., *tiáo* 'CLF:SLENDER', according to Frankowsky and Ke (2016: 63), and kind and event classifiers, as discussed in § 3.2.2. Most measure words can be used with more than one noun referent and be followed by *duō* 'more'. As shown in example (37), the measure word *xiāng* 'MENS:BOX' is used to denote quantity, and thus, it can be used to refer to a set of referents as in (b), and be followed by *duō* 'more' to specify additional quantity as in (c).

(37) The use of measure words

a. *yī*   *xiāng*   *píngguǒ*
    one   MENS:BOX   apple
    'one box of apples'

b. *yī*   *xiāng*   *píngguǒ*   *hé*   *lí*
    one   MENS:BOX   apple   and   pear
    'one box of apples and pears'

c. *yī*   *xiāng*   *duō*   *píngguǒ*
    one   MENS:BOX   more   apple
    'more than one box of apples'

In contrast, numeral classifiers can only be used with one referent or be followed by *duō* 'more' based on the specific properties they denote. As shown in example (38), the numeral classifier *bān* 'CLF:SHIFT' cannot be replaced with *gè*, as they have a different meaning in the two phrases in (a) and (b). It can be differentiated from measure words by being

tested with the cooccurrence with more than one noun referent as in (c) or with the element *duō* 'more' as in (d).

(38) The use of specific numeral classifiers
    a. *yī*   *bān*   *fēijī*
       one   CLF:SHIFT   plane
       'one flight'
    b. *yī*   *gè*   *fēijī*
       one   CLF:GENERAL   plane
       'one plane'
    c. **yī*   *bān*   *fēijī*   *hé*   *huōchē*
       one   CLF:SHIFT   plane   and   train
       'one plane and train'
    d. **yī*   *bān*   *duō*   *fēijī*
       one   CLF:SHIFT   more   plane
       'one and a bit more plane'

As to *de*-insertion, the genitive particle *de* can only be inserted between measure words and nouns, according to Wang (1994: 27-36). Later studies showed that *de* can also follow numeral classifiers, typically when the numeral they cooccur with is not *yī* 'one' or when the numeral classifier is pre-modified by an adjective. As shown in example (39), *de* is found between the numeral classifiers and nouns; it would not appear in (a) if the numeral is *yī* 'one' nor in (b) when the adjective is absent.

(39) The insertion of *de* in classifier phrases (Tang 2005: 445-446)
    a. *yībǎi*   *zhāng*   *de*   *fāng*   *zhuōzǐ*
       One-hundred   CLF:SPREADING.OPEN/FLAT   MOD   square   table
       'one hundred square tables'
    b. *xiǎo*   *tiáo*   *de*   *yú*
       small   CLF:SLENDER   MOD   fish
       'small fish'

De-insertion between numeral classifiers and nouns has been interpreted in terms of 'information weight' (Tang 2005: 444) and 'computational complexity' (Her and Hsieh 2010: 540). According to Tang (2005: 444), the higher the number, the 'heavier' the information. Her and Hsieh (2010: 540) argue that the 'weight' involved is not only related to the absolute value of the number but also the 'complexity' of information expressed in numerals and numeral classifiers. Therefore, *de*-insertion is likely to be induced when the numeral is smaller or larger than one or when the numeral classifier is pre-modified with an adjective. Therefore, the insertion of the genitive particle *de* is not a defining feature

of numeral classifiers or measure words but is motivated by the complexity of the information expressed in the quantification structure.

The criteria concerning omission in listing items and use of adjectives as pre-modifiers can be related to formal distinctions between numeral classifiers and measure words. As formal criteria are regarded necessary when differentiating one category from another, according to Corbett (1991: 147), the following part will focus on the formal aspects of the two criteria.

According to Wang (1994: 27-36), numeral classifiers can be omitted in listing items without changing the meanings of numeral noun phrases. However, measure words cannot be omitted in noun phrases. This criterium was attested by Her (2012: 4) in the comparison of sentences with and without the presence of numeral classifiers and measure words. As shown in example (40), numeral classifiers are present between numerals and nouns in sentence (a) but absent in (b). However, the two sentences mean the same.

(40) The omission of numeral classifiers in numeral noun phrases (Her 2012: 4)

a.
wǔ	*zhāng*		*bǐng*	*èr*	*tiáo*	*yú*
five	CLF:SPREADING.OPEN/FLAT		loaf	two	CLF:SLENDER	fish
wèi bǎo	*wǔqiān*		*gè*		*rén*	
feed.full	five-thousand		CLF:GENERAL		person	

'5000 people were fed by 5 loaves and 2 fish.'

b.
wǔ	*bǐng*	*èr*	*yú*
five	loaf	two	fish
wèibǎo	*wǔqiān*	*rén*	
feed.full	five-thousand	person	

'5000 people were fed by 5 loaves and 2 fish.'

A different scenario is shown in example (41). With the measure words inserted between numerals and nouns in (a), the sentence has a different meaning from the sentence in (b) without measure words.

(41) The omission of measure words in numeral noun phrases (Her 2012: 4)

a.
wǔ	*lán*		*bǐng*	*èr*	*xiāng*	*yú*
five	MENS:BASKET		loaf	two	MENS:BOX	fish
wèibǎo	*wǔqiān*		*zǔ*		*rén*	
feed.full	five-thousand		MENS:GROUP		person	

'5000 groups of people were fed by 5 baskets of loaves and 2 boxes of fish.'

b. *wǔ*   *bǐng*   *èr*   *yú*
    five   loaf   two   fish
    wèibǎo   *wǔqiān*   *rén*
    feed.full five-thousand   person
    '5000 people were fed by 5 loaves and 2 fish.'

Her (2012: 4) accounted for the formal differences between numeral classifiers and measure words in terms of the scope of numerals. According to him, numerals used with numeral classifiers take scope over nouns, and thus, numeral classifiers can be omitted in listing items or when stylistically required. In contrast, measure words block the scope of numerals to quantify nouns, and therefore, the omission of measure words can lead to changes in the meanings of the phrase.

As to adjectives used as pre-modifiers, Wang (1994: 27-36) proposed that adjectives can be used to modify measure words instead of numeral classifiers. However, later studies showed that adjectives can precede both measure words and numeral classifiers (e.g., Tang 2005; Her and Hsieh 2010). Her (2012: 4-9) distinguished between numeral classifiers and measure words based on the distribution of adjectives in numeral noun phrases. According to him, the distribution of adjectives can affect the syntactic interpretation of numeral classifiers and measure words. As shown in example (42), phrases (a) and (b) both mean 'one big apple', regardless of whether the adjective *dà* 'big' is placed before or after the numeral classifier *kē* 'CLF:ROUNDISH'. However, phrases (c) and (d) have different meanings due to the shift of the position of the adjective *dà* 'big': the numeral noun phrase in (c) means 'one big box of apples', while the numeral noun phrase in (d) means 'one box of big apples'.

(42) The distinction of numeral classifiers and measure words with adjective modifiers (Her 2012: 5)
    a. *yī*   *dà*   *kē*   *píngguo*
        one   big   CLF:ROUNDISH   apple
        'one big apple'
    b. *yī*   *kē*   *dà*   *píngguo*
        one   CLF:ROUNDISH   big   apple
        'one big apple'
    c. *yī*   *dà*   *xiāng*   *píngguo*
        one   big   MENS:BOX   apple
        'one big box of apples'
    d. *yī*   *xiāng*   *dà*   *píngguo*
        one   MENS:BOX   big   apple
        'one box of big apples'

63

The differences between the two pairs of numeral noun phrases in example (42) can be attributed to the syntactic features of numeral classifiers and measure words. According to Her (2012: 4-7), pre-classifier adjectives can also be modifiers of nouns, as both numeral classifiers and nouns are within the scope of the adjective modification. The structure of numeral noun phrases with pre-classifier adjectives is illustrated in Fig. 9. As shown in the figure, the pre-classifier adjective (ADJ) on the left branch is mapped to the position of ADJ' as the modifier of the noun. In such a structure, the distribution of adjectives does not affect the interpretation of the meanings of numeral noun phrases with numeral classifiers.

NUM	ADJ	CLF	ADJ'	N
yī	*dà*	*kē*	*(dà)*	*píngguo*
one	big	CLF:ROUNDISH		apple

'one big apple'
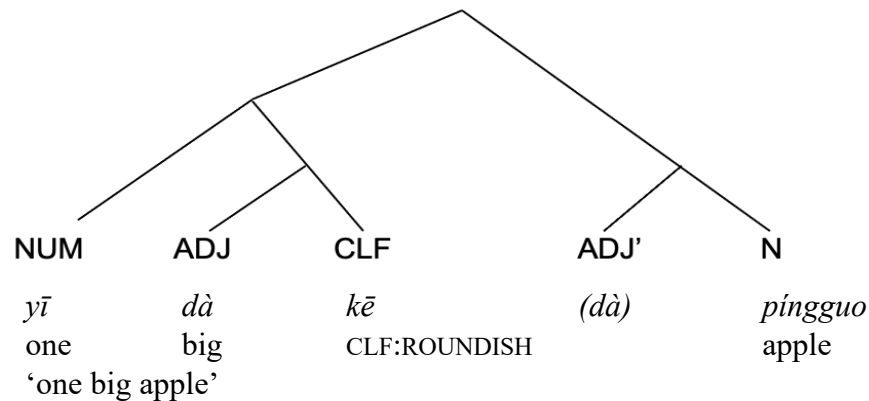
Fig. 9. The structure of Chinese numeral noun phrases with pre-classifier adjective (Her 2012: 4-7).

However, the distribution of adjectives can influence the interpretation of numeral noun phrases with measure words, as measure words block the scope of preceding adjectives to nouns. As shown in Fig. 10, the two adjectives are modifiers of measure words and nouns, respectively. The switch of the two adjectives can lead to different interpretations of numeral noun phrases.

NUM	ADJ1	MENS	ADJ2	N
yī	*xiǎo*	*xiāng*	*dà*	*píngguo*
one	small	MENS:BOX	big	apple

'one small box of big apples'
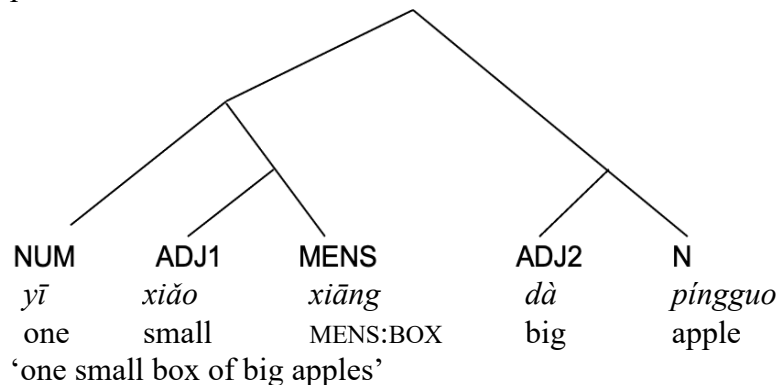
Fig. 10. The structure of Chinese numeral noun phrases with measure words and adjective modifiers.

In conclusion, Chinese numeral classifiers can be distinguished from measure words based on semantic and formal criteria. Most measure words can be differentiated from numeral classifiers based on the criterion of *ge*-substitution. Some specific classifiers denoting shape, kind or event can be distinguished from measure words, if they cannot cooccur with more than one noun referent or with the particle *duō* 'more'. Numeral classifiers and measure words can also be differentiated based on formal criteria concerning their omission or the distribution of adjectives in noun phrases. The deletion of numeral classifiers does not lead to changes in the meanings of noun phrases, while measure words cannot be omitted in noun phrases. Furthermore, the distribution of adjectives does not influence the interpretation of the meanings of classifier phrases, while noun phrases with measure words may differ in their meaning depending on whether adjectives are used to modify measure words or nouns.

3.3. Functions of Chinese numeral classifiers

As shown in §2.4.2, classifier systems have two major functions, i.e., semantic and pragmatic. While previous studies on the functions of Chinese numeral classifiers tended to focus on one aspect, typically individuation (e.g., Cheng and Sybesma 1998, 1999) and definiteness (e.g., Chen 2003; Li and Bisang 2012), the classifiers share most functions of other classifier systems discussed above. These functions will be discussed in §3.3.1 and §3.3.2, following the framework proposed by Contini-Morava and Kilarski (2013).

3.3.1. Semantic functions

According to Contini-Morava and Kilarski (2013: 268-279), nominal classification markers, in general, have four semantic functions. They can be used to expand the lexicon, differentiate referents by specifying their properties, individuate nouns by signalling information related to number and ascribe properties to referents, as discussed in §2.3.4. The uses of Chinese numeral classifiers can be classified in terms of three among these functions, i.e., "differentiating referents", "individuation", and "ascribing properties to referents".

Like other classifier systems, numeral classifiers in Chinese are not typically used to create new lexical items. However, some numeral classifiers have developed into derivational suffixes. For example, the morpheme *yuán* 'member', equivalent to the suffix -*er* in English, is frequently used to form words related to new occupations or professions, as in *yǎnyuán* 'act-member, actor' and *kuàidìyuán* 'express delivery-member, courier'. However, it used to be a common classifier used with nouns for people with a high social status in Ancient Chinese in Song and Yuan dynasties (1127-1368) (Ma 2015: 146). The use of classifiers as suffixes will be discussed further in §3.4.2.

Among the more typical semantic functions, Chinese numeral classifiers can differentiate noun referents. As the choice of numeral classifiers is semantically motivated, the features of noun referents can be differentiated when the same noun is used with different classifiers. Noun referents can be distinguished typically based on shape. In example (43), the two numeral classifiers *wān* 'CLF:CURVE' and *lún* 'CLF:WHEEL' distinguish between the different shapes of a curve-like crescent in (a) and a wheel-like full moon in (b).

(43) Chinese numeral classifiers used to differentiate noun referents based on shape
 a. *yī* *wān*   *yuèliang*
   one CLF:CURVE moon
  'one crescent'
 b. *yī* *lún*   *yuèliang*
   one CLF:WHEEL moon
  'one full moon'

Chinese numeral classifiers can also differentiate noun referents based on other features, such as an entity in opposition to a kind or an event. In example (44), three numeral classifiers are used with the noun *fēijī* 'plane': in (a) *jià* 'CLF:FRAME' indicates that the referent is a mechanism with supporting structures, in (b) *zhǒng* 'CLF:KIND, GENERAL' refers to a particular kind of plane, while in (c) *bān* 'CLF:SHIFT' is used with reference to a flight as an event. The use of the three different types of numeral classifiers to differentiate referents will be further analysed in Chapter 6.

(44) Chinese numeral classifiers used to differentiate noun referents based on entity in opposition to a kind or an event
 a. *yī* *jià*   *fēijī*
   one CLF:FRAME plane
  'one plane'

b. *yī*   *zhǒng*   *fēijī*
    one   CLF:KIND, GENERAL   plane
    'one kind of plane'
c. *yī*   *bān*   *fēijī*
    one   CLF:SHIFT   plane
    'one flight'

Secondly, Chinese numeral classifiers can be used to individuate nouns for the purpose of quantification.[8] The function of individuation is closely related to the unspecified nature of Chinese nouns, as discussed in §3.2.1. While the views vary on the nature of nouns in numeral classifier languages, it is generally agreed that Chinese numeral classifiers are used to create a semantic unit for the quantification of nouns. As discussed in §2.4.2, numeral classifier systems have been interpreted as complementary to plural marking in non-classifier languages (Borer 2005: 92-96), and numeral classifiers and plural markers, including -*men* in Chinese and -*s* in English, are unified under the same grammatical category (e.g., T'sou 1976; Borer 2005; Cowper and Hall 2012; Doetjes 2011; Her 2012; Mathieu 2012). According to Her (2012), both plural suffixes and numeral classifiers can function as multiplicands from a mathematical perspective. While Chinese numeral classifiers are generally obligatory in numeral noun phrases and can be omitted only for some syntactic considerations, the plural suffix -*s* in English is omitted and nouns are shown in singular form when the value of the multiplicand is 1.

While numeral classifiers and plural markers tend to be regarded as mutually exclusive (Greenberg 1972; Sanches and Slobin 1973; Tang and Her 2019), data from corpora and the Internet show that Chinese numeral classifiers and the plural marker -*men* do cooccur (Her and Chen 2013: 41). As shown in example (45), both the numeral classifier *wèi* 'CLF:INDIVIDUAL, RESPECT' and the plural marker -*men* occur with the noun *lǎoshī* 'teacher'.

(45) The cooccurrence of Chinese numeral classifiers and the plural marker on nouns (Her and Chen 2013: 41)

    san   *wèi*   *lǎoshī-men*
    three   CLF:INDIVIDUAL, RESPECT   teacher-PL
    'the three teachers'

---

[8] This function has been discussed by Cheng and Sybesma (1998, 1999), Ding ([1961] 1999), and Li et al. (2008).

Such constructions were interpreted by Her and Chen (2013: 41-42) as "double-headed", with the numeral and numeral classifier expressing definiteness and the plural marker *-men* providing a plural reading. What should be noted is that the cooccurrence of numeral classifiers and the plural marker *-men* used to be ungrammatical and is still uncommon in formal Chinese, and thus it cannot be regarded as one of the main features of modern Chinese. While numeral classifiers may cooccur with plural markers in Chinese, they should be regarded as definite determiners. The use of numeral classifiers for definiteness will be further discussed in §3.3.2.

Chinese numeral classifiers can also be used to ascribe properties to noun referents. Typical examples involve the use of *wèi* 'CLF:INDIVIDUAL, RESPECT' and *zūn* 'CLF:RESPECT', with the former used with nouns for people with a high social status, while the latter is used to show the high value of a statue or an artistic work. As illustrated in example (46), *wèi* 'CLF:INDIVIDUAL, RESPECT' in (a) shows that the noun *jiàoshòu* 'professor' refers to a respectable person, while *zūn* 'CLF:RESPECT' in (b) shows that the vase is valuable, and in (c) it shows that the statue of the Buddha should be respected or worshipped.

(46) Chinese numeral classifiers used to ascribe properties of noun referents
    a. *yī*   *wèi*   *jiàoshòu*
        one   CLF:INDIVIDUAL, RESPECT   professor
        'one professor'
    b. *yī*   *zūn*   *huāpíng*
        one   CLF:RESPECT   vase
        'one vase'
    c. *yī*   *zūn*   *fóxiàng*
        one   CLF:RESPECT   statue.of.the.Buddha
        'one statue of the Buddha'

Numeral classifiers can thus be used to show the attitude of a speaker to the addressee. For example, while *wèi* 'CLF:INDIVIDUAL, RESPECT' shows the social status as discussed above, it also expresses respect toward the professor. What should be noted is that the noun *jiàoshòu* 'professor' does not have to occur with the numeral classifier *wèi* 'CLF:INDIVIDUAL, RESPECT'. Instead, it can also be used with the general classifier *gè* 'CLF:GENERAL', especially when a speaker tries to be objective and self-effacing. However, if the professor in question is addressed or referred to, the replacement of *wèi* 'CLF:INDIVIDUAL, RESPECT' with the general classifier *gè* 'CLF:GENERAL' may imply contempt or irony, because the general classifier *gè* 'CLF:GENERAL' does not convey sufficient respect. In such

cases, the general classifier is used to degrade the social status of the referent and expresses contempt or irony.

The function involving ascribing properties to referents can restrict the choice of nouns. While *wèi* 'CLF:INDIVIDUAL, RESPECT' denotes humanness as mentioned in §3.2.2, it is not used with reference to people who commit crimes. Such nouns as *zuìfàn* 'criminal' and *qiángdào* 'robber' are generally excluded from among the nouns used with *wèi* 'CLF:INDIVIDUAL, RESPECT'. The numeral classifier *wèi* 'CLF:INDIVIDUAL, RESPECT' is also more likely to cooccur with nouns referring to the elders. Therefore, it is more likely to be used with such kinship terms as *zǔfù* 'grandfather', *mǔqīn* 'mother', and *jiějie* 'elder sister', and less likely to be used with *sūnzi* 'grandson', *nǚér* 'daughter', and *mèimei* 'younger sister'.

In conclusion, Chinese numeral classifiers can be used to differentiate referents, individuate nouns and express affective meanings. These semantic functions are shared by classifiers in general, except that numeral classifiers in Chinese can also be divided into entity, kind and event classifiers.

3.3.2. Discourse functions

The discourse functions of Chinese numeral classifiers involve reference identification, reference management, and re-presentation of referents, which are shared by all types of classifiers as shown by Contini-Morava and Kilarski (2013: 279-291).

First, numeral classifiers in Chinese can be used to identify referents in discourse. As indicated in example (47), the numeral classifier *zhī* 'CLF:BRANCH-LIKE' in the second clause is used anaphorically to refer to the pens mentioned in the first clause.[9]

(47) Anaphoric use of Chinese numeral classifiers

Tā	*yòu*	*ná*	*le*	*lìngwài*	*liǎng*	*sān*	*zhī*	*bǐ*
he	again	take	PRT	other	two	three	CLF:BRANCH-LIKE	pen

lái	*kàn,*	*měiyī*	*zhī*	*dōu*	*shì*	*kěyǐ*	*xiě*	*de*。
come	see	each	CLF:BRANCH-LIKE	all	be	can	write	MOD

'He took another two or three pens and found that each of them could write.'

---

[9] Examples in this section are collected on Google or Baidu unless otherwise indicated.

The anaphoric use is related to the use for deixis to indicate a referent that is evident in the context. The general classifier *gè* 'CLF:GENERAL' is the most typical one to be used with the demonstratives *zhè* 'this' and *nà* 'that' while pointing at a referent obvious to both parties in a conversation. It can indicate the referents even when the related nouns have not been explicitly mentioned. For instance, a customer can buy what they need in in a shop by using the general classifier *gè* in such sentences as in example (48) together with some gestures pointing at the merchandise.

(48) Deictic use of Chinese numeral classifiers

wǒ	*yào*	*mǎi*	*zhè*	*gè.*
I	want	buy	this	CLF:GENERAL

'I want to buy this.'

The function is also related to the use of numeral classifiers for disambiguation among previously mentioned referents, although in Chinese nouns are preferred for disambiguation between antecedents. As shown in example (49), a mother and a child use numeral classifiers to distinguish between two referents in their conversation. In (a) *běn* 'CLF:BOOK' refers to a loose-leaf notebook, while *zhāng* 'CLF:SPPREADING OPEN/FLAT' refers to a piece of paper. In (b) the child uses the numeral classifier *zhāng* 'CLF:SPPREADING OPEN/FLAT' to refer to paper even though the noun *zhǐ* 'paper' is not explicitly mentioned.

(49) Chinese numeral classifiers used for disambiguation

a.
Nǐ	*yào*	*yī*	*běn*	*hái*
you	wang	one	CLF:BOOK	or

zhǐshì	*yī*	*zhāng?*
just	one	CLF:SPREADING OPEN/FLAT

'Do you need the (loose-leaf) notebook or just a piece of (filler) paper?

b.
Yī	*zhāng*。
one	CLF:SPREADING OPEN/FLAT

'A piece of (filler) paper.'

The second discourse function of Chinese numeral classifiers involves reference management, which is related to definiteness, referentiality and topicality (Contini-Morava and Kilarski 2013: 283-284). According to Erbaugh (1986: 408), specific classifiers are more likely to be used to mark the first mention of a new referent, but they tend to be replaced by general classifiers in subsequent mentions. As illustrated in example (50), *jiǎotàchē*

'bicycle' is first used with the specific classifier *liàng* 'CLF:VEHICLE', and the general classifier *gè* in the subsequent sentence.

(50) Use of the general classifier for reference management in Chinese (Erbaugh 1986: 408) (glosses modified)

cóng	*nàibiān*	*guòlái*	*yī*	*gè*	*xiǎo*	*háizǐ,*
from	there	over.come	one	CLF:GENERAL	small	child
uh,	*.. qí,*	*qí,*	*qí*	*zhe*	*yī*	***liàng***
uh,	ride	ride	ride	PROG	one	**CLF:VEHICLE**
jiǎotàchē	*uh*	*shì*	*yī*	***gè***	*hěn*	*kěài*
bicycle	uh	be	one	**CLF:GENERAL**	very	cute
de	*xiǎo*	*de*	*jiǎo tà chē*			
MOD	little	MOD	bicycle			

'from over there comes a child, uh, ride…, ride, riding a (CLF:VEHICLE) bicycle, uh, (it) is a (CLF:GENERAL) very cute little bicycle.'

The presence of Chinese numeral classifiers can indicate definiteness and referentiality. Therefore, they are usually compared to determiners in non-classifier languages (Cheng and Sybesma 2012a). [10] As shown in example (51), numeral classifiers refer to indefinite and unspecific referents in (a) and (b), no matter whether the numeral *yī* 'one' is present or not. In (c), the numeral classifier still has an indefinite reading, but should be interpreted as specific and refers to the book that has been bought.

(51) Chinese numeral classifiers used to express indefiniteness

a.

wǒ	*xiǎng*	*mǎi*	*běn*	*shū.*
I	want	buy	CLF:BOOK	book

'I'd like to buy a book.'

b.

wǒ	*xiǎng*	*mǎi*	*yī*	*běn*	*shū.*
I	want	buy	one	CLF:BOOK	book

'I'd like to buy a book.'

c.

wǒ	*mǎi*	*le*	*yī*	*běn*	*shū.*
I	buy	PRT	one	CLF:BOOK	book

'I bought a/one book.'

Chen (2003) also found that numeral classifiers occurring in the structure of [CLF+N] can be used to express definite reference, typically when they are introduced by the object marker *bǎ* or followed by proper nouns and kinship nouns. The general classifier *gè* 'CLF:GENERAL' is more likely to be used in such structures. In the example in (52), *gè*

---

[10] See also Cheng and Sybesma (1999, 2012b), and Chen (2003).

'CLF:GENERAL' indicates definite reference and should be regarded as a definite determiner.

(52) Chinese numeral classifiers used to express definiteness (Chen 2003: 1174-1180)

tā	*bǎ*	*gè*	*píbāo*	*diū*	*le*
He	BA	CLF:GENERAL	bag	lose	PRT

'He lost his/the bag.'

Chinese numeral classifiers also exhibit discourse patterns involving thematic salience. Such classifier phrases occur more frequently in either foregrounded or presentative clauses (independent affirmative clauses, rather than negative and interrogative clauses), as shown by Li (2000) based on a study of spoken and written narratives. Furthermore, Li (2000) finds that compared with bare nouns, classified nouns can be modified by more words. Thematic salience expressed by numeral classifiers is illustrated in example (53).

(53) Use of Chinese numeral classifiers in discourse (glosses modified) (Li 2000: 1121-1122)

Chuánshuō	*zài*	*hěn*	*gǔ*	*de*	
legend say	be	very	old	MOD	
shíhòu,	***yǒu***	***yī***	***gè***		***jiào***
time,	There.be	one	CLF:GENERAL		called
Yōudū	***de***	***dìfāng,***	*zhōngnián*	*bú*	*jiàn*
Youdu	MOD	place	all.year	not	see
tàiyáng,	*dàochù*	*yīpiàn*	*qīhēi。*		*Zài*
sun,	everywhere	all	pitch.dark.		In
nàr	***yǒu***	***yī***	***zuò***	***dà***	***hēi***
there	there.be	one	CLF:STAND	big	dark
shān,	*shān*	*shàng*	*zhù*	*zhe*	*xǔduō*
mountain,	mountain	top	live	PROG	many
kěpà	*de*	*guàishòu。*		*Nàxiē*	*guàishòu*
scary	MOD	monster.		those	monsters
jīngcháng	*xià*	*shān*	*wēihài*	*rénmen。*	***Yǒu***
often	descend	mountain	endanger	people.	There.be
yī	***gè***		***jùrén***	***jiào***	***Kuāfù,***
one	CLF:GENERAL		giant	named	Kuafu,
tā	*yòng*	*guǎizhàng*		*hé*	*guàishòu*
he	use	cane		with	monster
bódòu	*le*	*jiǔ*	*tiān*	*jiǔ*	*yè,*
fight	PRT	9	day	9	night,
zhōngyú	*bǎ*	*tā*	*dǎ*	*sǐ*	*le。*
finally	BA	them	beat	dead	PRT.

'Once upon a time, in a place called Youdu, people lived in darkness all year round. There was a big black mountain where many terrible beasts lived. The beasts often went out to harm people. There was a giant called Kuafu. He fought with the beasts with a stick for nine days and nine nights. Finally, he killed them all…'

In example (53), three referents are introduced by *yǒu* 'there.be' in presentative structures: Youdu, a big black mountain, and a giant named Kuāfù. Their salience can be shown in two aspects. First, the nouns in the three phrases cooccur with prenominal modifiers, as shown in the first phrase pre-modified by *jiào Yōudū de* (call Youdu MOD) 'called Youdu'. According to Li (2000: 1120), noun phrases with numeral classifiers tend to take more prenominal modifiers than those without numeral classifiers. Second, the referents are also referred to two to three times in the subsequent discourse. However, such salience does not necessarily imply thematic importance, as numeral classifiers may only be used to highlight their referents instead of showing their thematic significance (Li 2000: 1118). This point can also be shown in example (52) above, where the general classifier is introduced by the object marker *bǎ*. According to Chen (2003: 1178-1179), numeral classifiers introduced by *bǎ* can be regarded as definite determiners, even though their referents are not mentioned in the following discourse.

Finally, Chinese numeral classifiers can also be used to recategorize noun referents in discourse. By assigning different numeral classifiers to the same noun referent, numeral classifiers may present the referent from different perspectives. This function is illustrated in example (54).

(54) Re-presentation of referents by different numeral classifiers

"Guāngdāng"	*yī*	*shēng,*	*Nīnà*	*bú*	*xiǎoxīn*
Bang	one	sound	Nina	not	carefully
pèngdào	*le*	*ménkǒu*	*de*	*yī*	*gè*
knock.over	PRT	door	MOD	one	CLF:GENERAL
cíqì 。	*Nà*	*shì*	*yī*	*zūn*	
ceramic.ware	That	is	one	CLF:RESPECT	

huāpíng	*yàng*	*de*	*dōngxī,*	*wàibiǎo*	*kànshàngqù*
vase	like	MOD	thing,	appearance	look

pō	*yòuxiē*	*gǔsègǔxiāng*		*de*	*yùnwèi* 。
rather	Some	antique		MOD	charm

'With a bang, Nina accidentally knocked **a ceramic ware** over at the door. It was **a vase-like thing, which looked rather antique**.'

As shown in example (54), both nouns *cíqì* 'porcelain' and *dōngxī* 'thing' refer to the same vase that Nina knocked over. However, it is first referred to by the general classifier *gè* 'CLF:GENERAL' and then by a more specific classifier *zūn* 'CLF: RESPECT'. The use of a general classifier may indicate that Nina was merely conscious of knocking over something but nothing specific. However, a closer look at the vase made her realize that she had broken something fragile but very likely antique and valuable. It shows that the choice of numeral classifiers is determined by different perspectives on the referent rather than the nature of nouns, as shown in the use of the general classifier *gè* with the more specific noun *cíqì* 'ceramic ware', and the classifier *zūn* 'CLF: RESPECT' with the more general noun *dōngxī* 'thing'.

In summary, this section described the discourse functions of Chinese numeral classifiers involving reference identification, reference management and representation of referents. As shown in this section, the presence of Chinese numeral classifiers can be used to identify the referents even when the related noun is omitted. Furthermore, they can be used to express definiteness, referentiality and topical salience. Therefore, they tend to be compared to determiners in non-classifier languages. Lastly, Chinese numeral classifiers are used to recategorize noun referents from different perspectives.

3.4. Diachrony of Chinese numeral classifiers

While Chinese has a long history of the use of numeral classifiers, it is less clear as to when they originated and how they developed. The origin of Chinese numeral classifiers in the context of their syntactic and semantic features will be discussed in § 3.4.1. In §3.4.2, I will deal with the development of the system of numeral classifiers.

3.4.1. The origin of Chinese numeral classifiers

Several hypotheses have been proposed as to the origin of numeral classifiers in Classical Chinese. Some scholars believe Chinese numeral classifiers appeared in the Pre-Qin period (770-221 BC) or even earlier (e.g., Guan 1953; Ma 2015: 63), while Peyraube (1991) claimed that they were formed as late as in the Tang period (618-907). Most scholars

argue that Chinese numeral classifiers first appeared during the periods from Han (206 BC-220 AD) to Weijin (220-420 AD) (e.g., Liu 1965; Wang 1994; Wu 2014a).

Three types of quantification structures can be identified in recorded documents before the Pre-Qin period (770-221BC): [NUM+N], [N+NUM], and [N1+NUM+N2]. As indicated in example (55), the numeral *sān* 'three' is followed by the noun *rén* 'person' in (a). The sequence of noun and numeral is reversed in (b) with the noun *níu* 'ox/cow' followed by the numeral *yī* 'one'. In (c) and (d), the quantification structure is composed of three elements: the first noun as N1, the numeral as NUM, and the second noun as N2. Usually, N2 is a repeater of N1 in their semantics. For example, *rén* 'person' and *hǔbēn* 'warrior' both refer to human beings, while *pǐ* 'horse' refers to a single horse, or a pair of male and female horses as in the Classical Chinese expression *mǔ pìn weì pǐ* (male.horse female.horse make a.pair.of.horses) 'one male and one female make a pair of horses' (Ma 2015: 48).

(55) Quantification structures before the Pre-Qin period, i.e., c.1050-771 BC (Wu 2014a: 84-85)

   a.   NUM   N
          sān   *rén*
          three   person
          'three people'

   b.   N   NUM
          níu   *yī*
          cow   one
          'one cow'

   c.   N1   NUM   N2
          hǔbēn   *sānbǎi*   *rén*
          warrior   three-hundred   person
          'Three hundred warriors'

   d.   N1   NUM   N2
          mǎ   *sì*   *pǐ*
          horse   four   horse
          'four horses'

Among them, the quantification structures [NUM+N] and [N+NUM] are predominant before the pre-Qin period (770-221 BC), while [N1+NUM+N2] is viewed as the source of numeral classifiers (Wu 2014a).

The origin of N2 is ascribed to the simplification of syllable structure and the loss of bimoraic foot in Classical Chinese, according to Bu (2011b) and Feng (2012: 91). Classical Chinese has undergone a simplification of consonant clusters and a loss of

bimoraic foot, typically between the Pre-Qin Periods (770-221BC) and Wei-Jin Periods and Southern-Northern Dynasties (220-589) (cf. Feng 1997: 212-213; 2012: 91). As a result, one single syllable was not 'heavy' enough to form a minimal independent prosodic unit—a foot. Therefore, N2 formed a prosodic foot as well as a constituent with numerals, especially when the numeral concerned was monosyllabic. Feng (1997: 232) also proposed the Sentential Prosodic Rule, according to which "a sentence is acceptable if the last element of the last phrase is properly assigned a stress". In the case of [N1+NUM+N2] as an independent clause or a phrase, an additional syllable provided by N2 was necessary to make the numeral structure 'heavy' enough in terms of the prosodic foot.

Before the first century AD, [NUM+N2] was usually used as a predicate rather than a modifier of N1. As shown in example (56), the constituents *yī gè* 'one bamboo' and *èr méi* 'two trunks' are both predicates, so that [N1+NUM+N2] can be used independently as a clause or a sentence as in (a). Therefore, [N1+NUM+N2] cannot be regarded as a construction in this period, nor can N2 be regarded as a numeral classifier, which is characterized by being more grammaticalized in the construction of numeral noun phrases.

(56) The use of *méi* and *gè* in the structure [N1+NUM+N2] before the first century AD in Chinese (cf. Wu 2014b: 64; Ma 2015: 53)
    a. *bā*   *cùn*   *zhú*   *yī*   *gè*
       eight   cun   bamboo one   N2-bamboo
       'one eight-cun bamboo'
    b. *yǒu*   *báishù*   *èr*   *méi*
       have   cypress two   N2-trunk
       'have two cypresses'

Numeral classifiers originated in the first century AD in the period of Han Dynasties (206 BC -220 AD) with the appearance of general classifiers and the construction of [NUM+CLF+N] with the constituent [NUM+CLF] used as a modifier of the head noun. A defining feature of numeral classifiers involving semantic reduction or semantic generalization can be shown in the appearance of general classifiers. General classifiers in Chinese did not appear until the first century AD in the Han Dynasty period (206 BC-220 AD), according to Wu (2014b: 64-71). For example, *méi* 'trunk' could only cooccur with a very limited range of nouns referring to trees, e.g. *báishù* 'cypress' in (b) in example (56) above, before the first century (Wu 2014a: 65-69). After the first century during the

Han Dynasty period (206 BC-220 AD), *méi* 'CLF: GENERAL, ROUND.PIECE' began to be used with almost all nouns for objects and animals, and can be regarded as a general classifier between the first and the third century (Wu 2014b: 64-71). Similarly, *gè* 'bamboo', which could only cooccur with the noun *zhú* 'bamboo' around the first century, expanded its scope of collocations thereafter.[11] From the third century, it began to replace *méi* as the general classifier by being assigned to almost all nouns (Liu 1965: 83).

Along with the generalization of *méi* 'trunk' and *gè* 'bamboo', more words were used as N2 in the structure [N1+NUM+N2]. They include such words as *běn* 'book' and *fēng* 'seal' in the pre-Qin period (770-221 BC), and *tóu* 'head' and *tiáo* 'little branch, slender shape' in the period of the Han Dynasties (206 BC -220 AD) (Ma 2015: 51-73). Some of them were used as specific classifiers used with a very narrow scope of nouns. For example, *běn* 'book' was exclusively used with nouns for books. Others cooccurred with a wider range of nouns. For example, *tóu* 'head' was used with nouns for animals, e.g., *yáng* 'goat, sheep', *niú* 'cow, ox', *yú* 'fish', and *yīng* 'eagle, hawk'.

With the generalization of *méi* 'trunk' and *gè* 'bamboo' and the use of more words as N2, the structure [NUM+N2] can be regarded as a constituent in which N2 is grammaticalized and semantically correlated with N1. This constituent could not only be used to enumerate N1 but also show some semantic features of N1, and thus, it developed some features of modifiers of N1. As modifiers in Chinese tend to precede nouns, the constituent of [NUM+N2] moved to precede N1 and led to the appearance of the construction of [NUM+CLF+N] (Wu 2014b).

3.4.2. The development of Chinese numeral classifiers

Chinese numeral classifiers developed as a result of the grammaticalization of nouns, verbs, and adjectives. Nouns are the major source of classifiers. As mentioned above, the general classifier *gè* 'CLF:GENERAL' was originally a noun referring to bamboo. The numeral classifiers that developed from nouns can also be used to denote animacy, shape,

---

[11] The written forms of *gè*, 个, 個 or 箇, remain controversial. Wang ([1954] 1985) and Liu (1965) argue that 个 is a simplified form of 個 and 箇 referring to bamboo, while Hong (1961) and Li and Zhang (2009) argue that the three characters are of two different origins, with 箇 referring to bamboo, while 个, as a simplified from of 個, developed from 介 *jiè* 'uniqueness'. Here, I follow Wang ([1954] 1985), Liu (1965), and Wu (2014a: 67-70) and regard them as being of the same origin.

as well as size. For example, *tóu* 'head' and *wěi* 'tail' are nouns when used independently but are numeral classifiers when used with nouns referring to, e.g., animals in numeral noun phrases. The classifiers *tiáo* 'CLF:SLENDER' and *kē* 'CLF:ROUNDISH', originally nouns referring to a branch of a plant and a fruit, respectively, are used to denote shape or size of their referents. Numeral classifiers can also develop from verbs. Such numeral classifiers include *zhāng* 'CLF:SPREADING.OPEN/ FLAT' used with such nouns as *zhǐ* 'paper' and *fēng* 'CLF:SEALING, ENVELOP' used with such nouns as *xìn* 'letter'. They developed from the two verbs *zhāng* 'spread' and *fēng* 'SEAL', respectively. Adjectives are a relatively rare source of Chinese numeral classifiers. For example, *duǒ* 'CLF:FLOWER-LIKE' developed from an adjective originally used to denote the blooming flowers or dangling fruits (Wu 2014b: 102-103). It is now a numeral classifier cooccurring with nouns for flowers or clouds to denote their flower-like shape.

Some Chinese numeral classifiers have further grammaticalized into bound morphemes (Loke 1997). While some elements, such as *tóu* 'head' and *wěi* 'tail' as mentioned above, can be used as both independent words and numeral classifiers, others can only be used as bound morphemes as numeral classifiers or further develop into derivational suffixes. For example, *yuán* 'member', a frequently used numeral classifier in Classical Chinese, is rarely collocated with nouns in Modern Chinese. Instead, it is more frequently used in the classification structure of [NUM+CL] or attached to other morphemes as a suffix to denote identity, occupation or profession, as shown in *xuéyuán* 'learn.member, learner' and *yǎnyuán* 'act.member, actor' in example (57).

(57) *Yuán* 'member' used as a bound morpheme in Modern Chinese
    xuéyuán   learn.member, 'learner'
    yǎnyuán   act.member, 'actor'
    guānyuán   official.member, 'official'
    shopyuán   shop.member, 'shop assistant'
    kuàidìyuán   express.delivery.member, 'courier'

The development of Chinese numeral classifiers also involves processes of semantic extension. The use of some numeral classifiers has been extended to a wide range of nouns based on physical features of referents. For example, *liàng* 'CLF:VEHICLE, CAR', used originally with *chē* 'carriage' since the pre-Qin period (770-221 BC), as mentioned in §3.2.2, is now used with nouns referring to vehicles, carts, trolleys, as well as artefacts with

wheels. Likewise, such numeral classifiers as *tóu* 'head' and *wěi* 'tail' mentioned above also categorize noun referents based on their salient features.

Other numeral classifiers less salient in their semantics undergo the processes of metonymic and metaphorical extension. A metonymic model assumes that "a word or expression normally or strictly used of one thing is used of something physically or otherwise associated with it" (Matthews 1997: 224). In terms of the development of Chinese numeral classifiers, a metonymical extension involves the transfer of the semantics of numeral classifiers when they are extended in their cooccurrence from prototypical noun referents to more peripherical members based on physical or other associated properties (Lakoff 1986: 31-33). Examples include the semantic extension of *gè* from the original reference to bamboo to other plants, as well as the above examples, including *tiáo* 'CLF:SLENDER', *kē* 'CLF:ROUNDISH', *zhāng* 'CLF: SPREADING OPEN/ FLAT', and *duǒ* 'CLF:FLOWER-LIKE'. For example, the meaning of *tiáo* 'CLF:SLENDER' has been extended from 'a long and slender branch of a plant' to the shape of being long and slender, as with such nouns as *lù* 'road', *gǒu* 'dog', and *hé* 'river', based on their physical properties.

Metaphor is regarded as a device that allows speakers to understand or experience an abstract concept in terms of a more concrete one (Lakoff 1986:31). The metaphorical extension of Chinese numeral classifiers can be shown when a numeral classifier denoting shape or other salient physical properties is extended to more abstract referents. A typical example is *tiáo* 'CLF:SLENDER'. As mentioned above, it has developed from a noun used with reference to a plant branch to a shape classifier by way of a metonymic extension. Its meaning has also metaphorically extended to allow it to cooccur with such abstract nouns as *xīnwén* 'news', *yìjiàn* 'advice, suggestion', *xiāoxī* 'message', and *lǐyóu* 'reason', based on their associated conceptual properties of being long and slender. Other numeral classifiers metaphorically extended to abstract concepts include *mén* 'CLF:GATE, BRANCH', as with *zhīshi* 'knowledge', and *dào* 'CLF:COURSE-LIKE, PATH-LIKE', as with *nántí* 'problem, challenge' and *nánguān* 'difficulty, barrier'.

In conclusion, Chinese numeral classifiers appeared in the first century AD in the period of Han Dynasties (206 BC -220 AD). They originated as a result of phonological changes in Classical Chinese and developed into obligatory elements in numeral noun phrases. The development of Chinese numeral classifiers involved the generalization of *méi* 'trunk' and *gè* 'bamboo' and the grammaticalization of other more specific classifiers in the construction of [N1+NUM+N2] or later [NUM+CLF+N], with the constituent of

[NUM+N2/CLF] used as prenominal modifiers. Like most classifiers, Chinese numeral classifiers are grammaticalized typically from nouns and more rarely from verbs and adjectives. Finally, the development of Chinese numeral classifiers also involved semantic extensions based on physical properties, and metonymy and metaphor.

3.5. Overview of the study on Chinese numeral classifiers
3.5.1. Traditional approaches to Chinese numeral classifiers

There is a considerable range of approaches to Chinese numeral classifiers in terms of their estimated number, typologies, and terminology in different linguistic traditions. This section will compare the different approaches to Chinese numeral classifiers in different linguistic traditions in the context of systems of quantification and nominal classification.

The estimated number of Chinese numeral classifiers varies in their descriptions from 22 (Erbaugh 1986), to 126 (Gao and Malt 2009), and to 427 (Huang and Ahrens 2003). Different approaches are also found in their typologies, ranging from just one type of numeral classifiers (Her 2017), two types of sortal and mensural classifiers in the Western tradition (Aikhenvald 2017: 374), to as many as 9 types in Chinese tradition, e.g., classifiers or individual measure, classifiers specially associated with V-O constructions, group measures, partitive measures, container measures, temporary measures, standard measures, quasi-measures, and measures for verbs (Chao 1968: 595). As regards the terminology, the most widely used term is 'numeral classifier' in the Western tradition (e.g., Aikhenvald 2000: 98-121; Bisang 2014), while other terms are also adopted by different scholars in different traditions, e.g., sortal classifiers in distinction from mensural classifiers (e.g., Her and Hsieh 2010), 'noun classifier' (e.g., Erbaugh 1986; Tai and Wang 1990; Gao 2010), 'individual classifier' (e.g., Gao and Malt 2009; Shao 2015), 'count-classifier' (e.g., Chien et al. 2003) and 'count-noun classifier' (e.g., Zhang 2007). Chinese numeral classifiers have also been referred to as quantifiers/measure words (e.g., Meng and Li 2011; Wang 2018), including individual quantifiers (e.g., Ma 2011), single quantifiers (e.g., He 2011), individual measure (e.g., Niu 2010) and individual measure words (e.g., Zhang and Xiao 2012; Liu 2016). This wide range of terms can be attributed to the different approaches to Chinese numeral classifiers in different linguistic traditions in

China and Western countries, as well as the different approaches to their place in both systems of quantifiers and nominal classification.

Chinese numeral classifiers have been approached differently in traditional Chinese linguistics and Western linguistics. In traditional Chinese linguistics, they were included in systems of quantifiers (or measure words). Quantifiers were first treated as a subtype of adjectives and regarded as elements used with numerals for the enumeration of their noun referents (Ma [1898] 2010: 121). The term *liàngcí*, literally 'quantifiers' or 'measure words' was first proposed by Li ([1924] 1992: 164-165), and in the next several decades, quantifiers were regarded as a subtype of nouns (Wang [1956] 2004: 272) or as general or 'attendant' nouns to denote quantities (Gao 1948; Liu 1965: 5). The term *gètí liàngcí*, literally 'individual quantifier/measures', was later proposed as a subtype of quantifiers or measure words by Ding ([1961] 1999: 174) and Chao (1968: 598), who defined them as units that individuate nouns and denote "shape, kind, or some other property associated with the noun". Therefore, *gètí liàngcí,* i.e., individual quantifiers/measures, in traditional Chinese linguistics are closest to numeral classifiers in Western linguistics in terms of individuation of nouns and their semantic correlation with nouns. Other than individual quantifiers/measures, other types of quantifiers that should also be treated as numeral classifiers include kind quantifiers or classifiers, and "classifiers specially associated with V-O constructions" by Chao (1968: 603-605) or (specialized) verbal quantifiers by such linguists as Ding ([1961] 1999: 178), which were later referred to as event classifiers by Huang and Ahrens (2003: 25-27).

The place of numeral classifiers in the system of quantifiers in traditional Chinese linguistics can be shown in Fig. 11. They are a type of nominal quantifiers, including individual quantifiers, specialized verbal quantifiers, and kind quantifiers. Measure words can be further subcategorized into six groups, with the first four denoting groups (e.g., *qún* 'MENS:CROWD'), parts or portions (e.g.,*fèn* 'MENS:PORTION' and *bùfèn* 'MENS:PART'), containers (e.g., *xiāng* 'MEN:BOX'), and arrangement (e.g., *duī* 'MEN:PILE' and *shù* 'MENS:BUNCH'), together with standard quantifiers denoting weight, distance, time, e.g. *dūn* 'MENS:TON', *gōnglǐ* 'MENS:KILOMETER', *xiǎoshí* 'MENS:HOUR', and temporary quantifiers borrowed from nouns, as in *yī shǒu níbā* (one hand mud, 'one handful of mud').
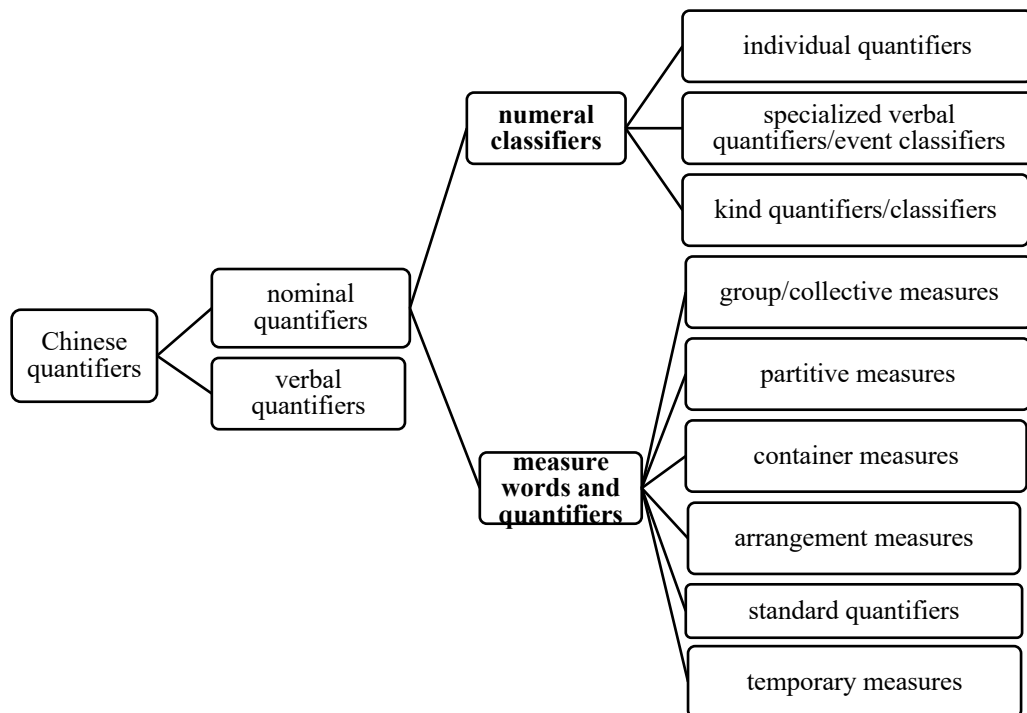
Fig. 11. The place of numeral classifiers in systems of quantifiers in traditional Chinese linguistics.[12]

Traditional Chinese linguists noticed the semantic correlation of individual quantifiers with their head nouns around the 1940s. For example, Gao (1948: 348) is among the very few Chinese linguists in this period who discovered that numeral classifiers can denote some semantic properties of nouns. However, such views have been largely marginalized in the next several decades, with the mainstream linguists focusing more on their syntactic features in quantification structures. Furthermore, the relationship between individual quantifiers and nouns was mainly regarded as random (Ding [1961] 1999: 174).

Very few studies have been made on Chinese numeral classifiers in traditional Western linguistics. Chinese numeral classifiers were first recorded by Western missionaries in such grammars as *Glossary of the Mandarin Language* (1703) by Francisco Varo (Coblin and Levi 2000) and *A Grammar of the Chinese language* by Robert Morrison (1782-1834) (Morrison 1815), or such dictionaries as *Chinese Latin Radicals Dictionary* (1694) (*Dictionarium Sinico Latinum*) and *Chinese Latin Pronunciation Dictionary (Dictionarium Sinico Latinum*, 1669) by Basillio Brollo (1648-1704) in the 16th-17th centuries (cf. Song 2014). Francisco Varo (1703) referred to them as 'particulas' (particles)

---

[12] Based on, e.g., Chao (1968: 595-631), Ding ([1961] 1999: 174-179), Huang and Ahrens (2003), and Zhang (2011: 4).

used with numerals or demonstratives (Coblin and Levi 2000). He also found that the general classifier *kó* (i.e., *gè*) could be applied to a wide range of nouns, while specific classifiers, e.g. *puèn* (i.e., *běn*), were more restricted in the collocations with nouns (Coblin and Levi 2000: 159). Thomas Francis Wade (1818-1895) noticed their correlation with nouns by defining Chinese numeral classifiers as "numerative nouns" or "associate (or attendant)" nouns used to refer to "form, use or an affinity" of their head nouns (cf. Song 2014). However, despite these individual works carried out by the western missionaries and diplomats in China, Chinese numeral classifiers have been regarded as pure linguistic forms in traditional Western linguistics. Furthermore, they were largely marginalized in the next two centuries in Western linguistics, as they were regarded as semantically redundant (e.g., Brinton 1885: 62) (see §2.6). Not until the 1960s did Chinese numeral classifiers attract attention in Western linguistics.

The above two approaches illustrate the place of Chinese numeral classifiers in traditional Chinese and Western linguistics. Traditional Chinese linguists focus more on their syntactic features, and therefore, classifiers have been regarded as one type of quantifiers. As all elements occurring with numerals or demonstratives in Chinese are regarded as quantifiers, Chinese quantifiers can be divided into two major types as shown in example (58), including verbal quantifiers occurring in verbal phrases, such as *yǎn* 'MENS:EYE' in (a), and nominal quantifiers including sortal classifiers or individual quantifiers such as *bě*n 'CLF:BOOK' in (b) and measure words such as *xiāng* 'MENS:BOX' in (c).

(58) Chinese verbal and nominal quantifiers
    a.   *kàn*   *yī*   *yǎn*
          see   one   MENS:EYE
          'have a look'
    b.   *yī*   *běn*   *shū*
          one   CLF:BOOK[13] book
          'a book'
    c.   *yī*   *xiāng*   *shū*
          one   MENS:BOX   book
          'one box of books'

---

[13] The English annotation of Chinese numeral classifiers is based on Gao and Malt (2009: 1171-1177) and my own understanding of them as a Chinese native speaker.

On the other hand, while some western missionaries and diplomats found some semantic features of Chinese numeral classifiers, they have been largely neglected in traditional Western linguistics.

3.5.2. Contemporary approaches to Chinese numeral classifiers

Contemporary studies on Chinese numeral classifiers are characterized by a convergence of the above-mentioned approaches in traditional Chinese linguistics and Western linguistics. With the systematic work conducted on nominal classification from 1970s in Western linguistics, Chinese numeral classifiers regained their attention in Western linguistics and were regarded as one type of classifiers attached or adjacent to numerals and closely correlated with their head nouns (e.g., Denny 1976; Allan 1977; Aikhenvald 2000) (see §2.4.1.1 and §2.6). Scholars began to describe Chinese numeral classifiers in the context of nominal classification. Chao (1968: 598) referred to those "interposed" elements in numeral noun phrases as "classifiers" or "individual measures". However, since the relationship between numeral classifiers and nouns had been regarded as random (Ding [1961] 1999: 174), little research was conducted on Chinese numeral classifiers from the perspective of nominal classification in the next two decades in Chinese linguistics, as mentioned in §3.5.1.

The 1980s can be regarded as the beginning of the studies on Chinese numeral classifiers in the contexts of both quantification and nominal classification. Erbaugh (1986: 402) distinguished Chinese numeral classifiers from measure words and defined them as obligatory elements with numeral and demonstratives and modifiers for concrete as well as abstract nouns. She examined the development of Chinese numeral classifiers in child language, compared it with their historical development, and found similar trends in both. In the 1990s, a series of studies were conducted to highlight the semantic features of numeral classifiers and their nominal categorization by Chinese scholars, including Tai and Wang (1990), Tai (1992, 1994) and Wang (1994).

Since the 1990s, Chinese numeral classifiers have been approached from several perspectives. These range from their syntactic features (e.g., Cheng and Sybesma 1999; Zhang 2012; Her 2017; Her and Tsai 2020) to their semantic properties (e.g., Huang and Ahrens 2003; Li 2011; Song 2017). Studies focused on their obligatoriness in relation to

the nature of Chinese nouns and pluralization (e.g., Chierchia 1998a; Cheng and Sybesma 1999; Borer 2005; Imai and Mazuka 2007; Her and Chen 2013; Wu and Her 2021) (see §3.2.1), the distinction between numeral classifiers and measure words (e.g., Wang 1994; Cheng and Sybesma 1998; Her and Hsieh 2010; Her 2012) (see §3.2.3), and definiteness, as shown by the debate between Cheng and Sybesma (1998, 1999) and Tang (2005), and later studies, e.g., by Cheng and Sybesma (2012a, 2012b), Li and Bisang (2012) and Li and Wu (2018). The obligatoriness and definiteness of Chinese numeral classifiers are closely related to language complexity, as illustrated by Bisang (2014) in his discussion of covert complexity. Studies have also been carried out on the functions of Chinese numeral classifiers. For example, Li (2000) focused on pragmatic functions, e.g., introducing foregrounded referents in presentative structures. Much attention has also been devoted to the diachrony of Chinese numeral classifiers, e.g., their origin and development (e.g., Peyraube 1991; Wang 1994; Ma 2015) and other more specific aspects, e.g., the role of phonological factors (Bu 2011b; Feng 2012), the development of classification constructions (Wu 2014a), and grammaticalization (e.g., Loke 1997; Jin and Chen 2002; Wang 2010; Meng and Li 2011). Comparative studies focused on Chinese numeral classifiers vs. classifiers in other Sinitic languages (e.g., Tai 1992; Cheng and Sybesma 2012b; Erbaugh 2013), other classifier languages in East Asia (e.g., Huang and Ochi 2014) or other parts of the world (e.g., Dong and Deng 2019; Her and Tang 2018).

Chinese numeral classifiers have also been the object of research in applied linguistics and psycholinguistics. As numeral classifiers are typical in Chinese while absent in most Indo-European languages, several studies dealt with the L2 acquisition of Chinese numeral classifiers (e.g., Hu 1993; Erbaugh 1986, 2006; Zhang and Lu 2013; Wang and Ren 2017; Crosthwaite et al. 2018). Psycholinguistic studies on the processing of classifiers have analysed both L1 and L2 speakers (e.g., Gao and Malt 2009; Bi et al. 2010; Srinivasan 2010; Hsu et al. 2014; Her et al. 2018).

3.6. Concluding remarks

This chapter has provided a sketch of Chinese numeral classifiers, including their syntactic and semantic features, semantic and discourse functions, and diachrony. Scholars working in different traditions have approached Chinese numeral classifiers from various

perspectives, including their properties, functions, as well as language acquisition and language processing. However, relatively few studies have been carried out on their semantic contribution, and on the comparison of Chinese numeral classifiers with their equivalent forms in English in terms of functionality. Therefore, the next chapters will address these issues based on corpus data.

Chapter 4: Methodology

4.1. Introduction

While numerous studies have been conducted on Chinese numeral classifiers, relatively few of them have been carried out on their functions and the degree of their equivalence with measure words and determiners in English, typically based on corpus and discourse data. In order to fill in the gap of research, both quantitative and qualitative studies were conducted. Quantitative studies dealt with semantic functions of Chinese numeral classifiers and their representation in English translations based on a corpus study. In order to examine the semantics of Chinese numeral classifiers, quantitative studies also examined their frequency and collocations. Qualitative studies investigated the discourse functions of Chinese numeral classifiers and their representation in English translation based on discourse data. Therefore, two corpora of numeral noun phrases were compiled for quantitative studies, and discourse data based on a novel were collected for qualitative studies.

Quantitative studies are discussed in §4.2, including the criteria and stages of data collection, sampling, annotation, cleaning, and methods of analysis, as well as two self-compiled corpora of numeral noun phrases. Section §4.3 deals with the methods for qualitative studies, including data collection, processing and analysis. Conclusions are given in §4.4.

4.2. Research methods for quantitative studies

A considerable number of studies have been conducted on the features and functions of Chinese numeral classifiers, as discussed in Chapter 3. However, relatively few studies have been based on corpus data. In this dissertation, two specialised corpora were compiled based on numeral noun phrases with and without pre-classifier adjectives. Specific methods concerning data collection, sampling, annotation, and cleaning are discussed as follows.

4.2.1. Data collection

Recent decades have witnessed the development of such types as parallel, balanced and specialized corpora in China. The widely acknowledged English-Chinese parallel corpora include Beijing Language and Culture Corpora (BCC) developed by Beijing Language and Culture Corpus Center (Xun et al. 2016) and the corpus of Centre for Chinese Linguistics of Peking University (CCL) (Zhan et al. 2003; Zhan et al. 2019). Other less well-known corpora include the BABEL English-Chinese Parallel Corpus (by Lancaster University) (Xiao 2004), the corpus of ChineseLDC (by Chinese Linguistic Data Consortium) (Ma 1999), and the TED English Chinese parallel corpus of speeches (collated by Beijing Foreign Studies University) (Xu 2019b; Yang 2021). While all the above parallel corpora are aligned, only some of them are tagged for part of speech (POS), including BCC, CCL, and BABEL. As to genre, the large-scale corpora BCC and CCL cover texts of various genres, while others are limited to one or several genres. For example, the texts collected in ChineseLDC are restricted to the genre of legal language, press releases and news items, while the TED English Chinese parallel corpus of speeches is limited to transcripts of TED talks. Regarding accessibility, ChineseLDC is accessible for commercial use, and other corpora are available online for research purpose for short period of time or for designated users. For example, the BCC English-Chinese parallel corpus was available for a short period of time in 2019 but has not been available since 2020, and CCL is only available for researchers on-site at Peking University. Finally, among all the above corpora, BCC is the only corpus that can be queried based on regular expressions.

Several balanced corpora have also been developed, including the frequently used Lancaster Corpus of Mandarin Chinese (LCMC) (McEnery and Xiao 2004) and the Academia Sinica Balanced Corpus of Modern Chinese (Sinica Corpus) (Huang and Chen 2010).[14] Specialized parallel corpora have also been developed for research for specific purposes. For example, the CLIPS corpus (Chinese Learners' Integrated Pear Stories corpus, by Beijing Foreign Studies University) is designed for language learning and acquisition (Xu 2019a), while corpora of different translation versions of such novels as *A Dream of Red Mansions* (Ren et al. 2010) and *Moment in Peking (*Zhang et al. 2011) are meant for the studies on corpus-based translation.

Among the above corpora, the BCC English-Chinese parallel corpus was chosen as the source corpus for the quantitative studies in this research project. First, as BCC is a large-scale corpus of a total size of about 10 billion tokens and covering various genres, it can be expected to provide a representative illustration of Chinese numeral classifiers. Second, BCC is tagged by part of speech and allows online queries based on regular expressions. Therefore, the corpus can be used to distinguish between numeral classifiers and other parts of speech and to query for numeral classifiers based on the constructions they occur in.

However, some limitations can also be found in the corpus. For example, both numeral classifiers and measure words are tagged by 'q' (quantifiers) in the corpus. Therefore, numeral classifiers need to be distinguished manually from measure words. Second, ungrammatical translations can also be found in the corpus. As shown in example (59), sentence in (a) in Chinese is an ungrammatical translation of the sentence in (b) in English.

---

[14] Other balanced corpora also include the corpus developed by the Institute of Applied linguistics of the State Language Commission (CNcorpus) (Jin et al. 2005; Xiao 2010, 2016), and the Texts of Recent Chinese corpora 2019 (ToRCH Corpora 2019*)* developed by Beijing Foreign Studies University (Li et al. 2022).

(59) Mistranslated sentence from BCC corpus (Xun et al. 2016)

  a. 一名在小，越来越多的德尔雷海滩，佛罗里达州，七城市沙质很多青少年锤击在炎热的中午太阳铲。

yī	*míng*	*zài*	*xiǎo,*	*yuèláiyuè*
one	CLF:IDENTITY	in	small,	more and more
duō	*de*	*Déěrléi*	*hǎitān,*	*Fóluólǐdá*	*zhōu,*
many	MOD	Delray	beach,	Florida	state,
qī	*chéngshì*	*shāzhì*	*hěnduō*	*qīngshàonián*
seven	city	sandy	many	teenager
chuíjī	*zài*	*yánrè*	*de*	*zhōngwǔ*
hammer	in	hot	MOD	mid-day
tàiyáng	*chǎn.*			
sun	shovel			

  b. *On a sandy lot in the small, growing city of Delray Beach, Florida, seven teenagers are hammering and shoveling in the hot mid-day sun.*

Among many grammatical mistakes in the Chinese sentence, the constituent *yī míng* 'one CLF:IDENTITY' at the beginning is followed by the preposition *zài* and the adjective *xiǎo* 'small', in an ordering which is ungrammatical in Chinese. Such sentences with grammatical mistakes typically involving numeral noun phrases were removed from the results. Third, the results of the queries are shown in sentences without POS-taggers on webpages and they cannot be directly downloaded. Therefore, the results have to be copied to spreadsheets and numeral noun phrases must be manually annotated. The data based on BCC were thus collected and processed in the following steps: online query, sampling, annotation, and data cleaning.

1) Query

In order to query for the sequences of [NUM+CLF/MENS] and [NUM+ADJ+CLF/MENS], two regular expressions were used: [一 q] and [一 a q].[15] The two regular expressions were chosen based on the following considerations. First, numeral noun phrases in Chinese are usually composed of three basic elements: numeral, numeral classifier or measure word, and noun, as shown above in §3.2.1. Second, the results of the query based on [NUM+CLF/MENS] include three types of classifier phrases, as shown in Table 7:

---

[15] The regular expressions in the BCC are based on Chinese characters, English words, and POS taggers. In these two regular expressions, the elements can be explained as follows: "一" (*yī* 'one'), "*q*" (numeral classifiers and measure words), and "*a*" (adjectives).

Table 7. Three types of classifier phrases without pre-classifier adjectives

Structures	Examples
[NUM+CLF/MENS]	*yī yuán* (one CLF:MEMBER), 'one member'
[NUM+CLF/MENS+N]	*yī gè dìfāng (*one CLF:GENERAL place), 'one place'
[NUM+CLF/MENS+MOD+N]	*yī gè tíng zhe de chē (*one CLF:GENERAL park PROG MOD car), 'one parked car'

However, the results do not include classifier phrases with pre-classifier adjectives. Therefore, the other query was made based on the second regular expression and the results based on [NUM+ADJ+CLF/MENS] include another three types of classifier phrases shown in Table 8.

Table 8. Three types of classifier phrases with pre-classifier adjectives

Structures	Examples
[NUM+ADJ+CLF/MENS]	*yī xiǎo kuài* (one small CLF:LUMP.LIKE), 'a small slice'
[NUM+ADJ+CLF/MENS+N]	*yī xiǎo kuài bù* (one small CLF:LUMP.LIKE cloth), 'a small piece of cloth'
[NUM+ADJ+CLF/MENS+MOD+N]	*yī xiǎo kuài shēnlánsè de bù* (one small CLF:LUMP.LIKE dark.blue MOD cloth), 'a small dark blue piece of cloth'

Third, numerals are limited to *yī* 'one', as it is the equivalent of the numeral *one* as well closest to the indefinite article in English, while other numerals in Chinese tend to be equivalent only to their corresponding numerals in English. As a result, more than 0.6 million pairs of Chinese-English sentences were derived based on the sequence of [NUM+CLF/MENS] and about 1,100 pairs of Chinese-English sentences were derived based on the sequence of [NUM+ADJ+CLF/MENS].

2) Sampling

Sampling was only made based on more than 0.6 million pairs of Chinese-English sentences with the sequence of [*yī* 'one'+ CLF/MENS]. In every 4,000 pairs of sentences, 50 of them were chosen as samples for the study. As a result, altogether 8,400 numeral noun phrases without adjectives were selected based on 7,700 pairs of sample sentences. As there were significantly fewer sentences in the results with the sequence [*yī* 'one'+ADJ+CLF/MENS], all 1,100 numeral noun phrases with pre-classifier adjectives were chosen based on 1,100 pairs of Chinese-English sentences.

3) Annotation

Numeral noun phrases in the sample sentences were manually annotated in terms of five categories in Chinese and their equivalents in English translations, as listed in Table 9.

Table 9. List of the categories of annotation

Chinese	English
numeral (NUM CN)	numeral, pronoun, and determiner (DET EN)
adjective (ADJ CN)	adjective (ADJ EN)
numeral classifier (CLF CN)	measure word (MENS EN)
modifier (MOD CN)	modifier1 (MOD1 EN)
	modifier2 (MOD2 EN)
noun (N CN)	noun (N EN)

As only one-word adjectives can precede numeral classifiers in Chinese, the elements pre-modifying numeral classifiers in Chinese and their equivalents in English were annotated as 'adjective'. In contrast, the elements modifying nouns were annotated as 'modifier', as there can be more words used to modify nouns. Modifiers in Chinese can be translated into equivalents occurring either before and after nouns in English, and therefore, the English modifiers were annotated as modifier 1 as premodifiers and modifier 2 as postmodifiers depending on their positions in relation to nouns.

The five pairs of categories were aligned except for the modifier 2 occurring after nouns in English, as illustrated in Fig. 12, based on two examples of *yī shǒu ōuwén de shī* 'one poem by Owen' in (a) and *yī cháng tiáo fánxīngmìbù de tiānkōng* 'one long strip of starry sky' in (b).

	NUM CN	DET EN	ADJ CN	ADJ EN	CLF CN	MENS EN	MOD CN		MOD1 EN	N CN	N EN	MOD2 EN
a.	*yī*	*a*	/	/	*shǒu*	/	*ōuwén*	*de*	/	*shī*	*poem*	*by Owen*
	one				CLF:SONG, POEM		*Owen*	MOD		poem		
	'one poem by Owen'											
b.	*yī*	one	*cháng*	/	*tiáo*	*strip of*	*fánxīngmìbù*	*de*	*starry*	*tiānkōng*	*sky*	/
	one		long		CLF:SLENDER		stars.densely. distributed	MOD		sky		
	'one stretch of mountains'											

Fig. 12. Sample alignment of the elements in numeral noun phrases in Chinese and English.

The five pairs of elements are not strict equivalents based on literal translations but rather aligned categories based on their parts of speech. For example, in (a), the numeral classifier *shǒu* 'CLF:SONG, POEM' is not directly translated into English, and therefore, its equivalent category was left blank. As to the nouns and modifiers in (a), the noun *shī* 'poem'

was aligned with its equivalent *poem*, and the equivalent modifiers *ōuwén de* (Owen MOD, 'Owen's') were translated into *by Owen* as modifier 2 in English, and modifier 1 was left blank. In (b), the adjective *cháng* 'long' is not directly translated, and therefore, its equivalent category in English was left blank.

Aside from the annotation of the above five categories, Chinese nouns were also annotated in terms of four semantic oppositions: count vs. mass, concrete vs. abstract, animate vs. inanimate, and human vs. nonhuman. The distinction between concrete and abstract nouns in Chinese can be made based on spatial or temporal grounds. Lyons (1977: 442-447) distinguished three types of nouns: a) physical objects located in a three-dimensional space, b) events and process distinguishable in time, and c) concepts non-observable either in space or time. In this study, Lyons' first type of nouns were regarded as concrete nouns, while the other two types were regarded as abstract nouns. Based on these distinctions, subtypes and examples of concrete and abstract nouns in the corpora were listed as follows. Concrete nouns whose referents can be physically identified include animate nouns for humans and nonhumans, and nouns for animals.

(1) [animate]: including nouns for human, animals, birds, insects, and fish. Two subcategories were distinguished in animate nouns, including:

    a. [human]: including nouns for humans, e.g., *rén* 'person or people', *fùqīn* 'father', and *jiàoshī* 'teacher'.

    b. [nonhuman]: including nouns for animals, birds, insects, and fish, e.g., *gǒu* 'dog', *niǎo* 'bird', *yú* 'fish', *chóngzi* 'worm', and *xìjūn* 'bacterium'.

(2) [inanimate]: including nouns for plants, body parts, buildings, instruments, commodities, means of transportation, natural entities, products, food, and nouns for groups and organizations, e.g., *shù* 'tree', *shǒu* 'hand', *dàolù* 'road', *gǎngkǒu* 'harbour', *jīqì* 'machine', *jiājù* 'furniture', *gāngbǐ* 'pen', *fēijī* 'plane', *xíngxīng* 'planet', *shānqiū* 'hill', *shēngyīn* 'sound', *kàngshēngsù* 'antibiotic', *dàngāo* 'cake', *chéngshì* 'city', and *fǎyuàn* 'court'.

Abstract nouns belong to the second and third types proposed by Lyons (1977: 442-447). They include nouns for observable events and processes, and those for non-observable facts and concepts. Examples are given as follows:

(1) nouns for observable events, changes and processes, e.g., *shìjiàn* 'event', *biànhuà* 'change', *gǎigé* 'reform', *jìnbù* 'progress', *diàochá* 'suvey,

investigation', *yùndòng* 'campaign, movement', *jùhuì* 'gathering', and *chēhuò* 'car accident';

(2) nouns for non-observable concepts, including nouns related to an idea (e.g., *xiǎngfǎ* 'idea, notion' and *xìnniàn* 'belief'), emotion (e.g., *jīngxǐ* 'surprise' and *yālì* 'pressure'), fact (e.g., *xiànxiàng* 'phenomenon' and *yōushì* 'advantage'), time, place and manner (e.g., *jiēduàn* 'stage', *fāngfǎ* 'manner', and *huánjìng* 'environment'), language and linguistic activities (e.g., *shùzì* 'number', *shēngmíng* 'statement', and *chéngnuò* 'promise'), science and education (e.g., *chéngxù* 'application' and *biānjíqì* 'editor'), value (e.g. *chéngběn* 'cost' and *cáichǎn* 'property'), and degree (e.g., *xiànzhì* 'limit', *jíbié* 'level').

Problematic cases were found in polysemous nouns and in some nouns referring to organizations. Polysemes can be both concrete and abstract. For example, *qiáoliáng* 'bridge' is concrete referring to a structure over rivers or roads, but abstract when it denotes a connection or contact between different things. Similarly, some nouns for social units and organizations, e.g., *jiātíng* 'family', *guójiā* 'country, nation', *shìjiè* 'world', and *shèhuì* 'society', can be concrete in terms of being perceptible or accessible in space and be abstract by being related to something conceptual in their functions, contents, activities, or events involved. These nouns were annotated as both concrete and abstract (±concrete).

While controversy remains as to whether there are countable nouns in Chinese, the distinction between countable vs. uncountable nouns can be made based on whether or not nominal referents can be separated in terms of physical and temporal bounds or internal features (Cheng and Sybesma 1999: 515; Huang and Ahrens 2003; Crystal 2008: 119; Zhang 2012: 11-13), as discussed in §3.2. Countable nouns were tested by being used with adjectives denoting size, dimension, and shape, typically *dà* 'big' and *xiǎo* 'small', as when nouns can be described in terms of these features, they can be regarded as separable and thus countable. In contrast, nouns that can only be modified by *yīdiǎndiǎn* 'a little bit', similar to *much* in English, were regarded as mass nouns in Chinese. As a result, most concrete nouns in the samples are countable (count), except for nouns for liquid and gas, e.g., *shuǐ* 'water' and *qì* 'gas', nouns for substances, e.g., *dànbáizhì* 'protein', *sèsù* 'pigment', *fángfǔjì* 'preservative' and *kàngshēngsù* 'antibiotic'. As regards abstract nouns, most nouns for events and processes are countable, while most nouns for emotions and feelings are uncountable (noncount). The remaining nouns were annotated as both countable and uncountable (±count), typically when they can be

modified by such adjectives as *dà* 'big' and *xiǎo* 'small', and the adverb *yīdiǎndiǎn* 'a little bit', as illustrated in example (60).

(60) Chinese nouns annotated by (±count)
  a. *hěn*   *dà*   *de*   *jīhuì*
       very   big   MOD   chance
       'very big chance'
  b. *yīdiǎndiǎn*   *jīhuì*
       a little bit   opportunity
       'a little bit of opportunity'

The noun *jīhuì* 'chance, opportunity' can be modified by both the adjective *hěndà de* 'very big MOD, very big', and the adverb *yīdiǎndiǎn* 'a little bit', and thus it was annotated as ±count. Similar words include *fāngfǎ* 'way, means, method', *gōngzuò* 'job, work', *gōngnéng* 'function, functionality', *huódòng* 'activity, movement', and *xiànxiàng* 'phenomenon, appearance'.

4) Data cleaning

Data cleaning was processed in three steps along with the sampling and annotation. First, repeated, misaligned or mistranslated sentences were excluded from the samples. Such mistranslations as shown in example (59) above were removed from the samples. Other items excluded from the results include misaligned sentences without equivalent translations of numeral noun phrases.

The second step was taken to exclude classifier phrases without nouns to examine the collocations of numeral classifiers and nouns. For example, the numeral classifier *yuán* 'CLF:MEMBER' occurred only in phrases without a following noun, as shown in example (61), while the traditional collocations of *yuán* 'CLF:MEMBER' with such human nouns as *měngjiàng* 'valiant general/person' and *gànjiàng* 'capable general/person' were not found in the corpus. Phrases as in the example below were removed from the results.

(61) Classifier phrases of *yuán* 'CLF:MEMBER' in BCC (Xun et al. 2016)

a. | ...*tā* | *shì* | *tàiyángxì* | *zhōng* | *de* | *yī* | *yuán*。 |
    |---------|-------|------------|---------|------|------|---------|
    | ...it | is | Solar System | among | MOD | one | CLF:MEMBER |

'...it's part of the Solar System.'

b. | *wǒ* | *yě* | *xiǎngyào* | *quèbǎo* | *měi* | *gè* |
    |----------|---------|------------|----------|----------|-------------|
    | I | also | want | ensure | every | CLF:GENERAL |
    | *dōu* | *gǎndào* | *zìjǐ* | *shì* | *tuánduì* | *de* |
    | all | feel | self | is | team | mod |
    | *rén* | *yī* | *yuán* ...... | | | |
    | person | one | CLF:MEMBER | | | |

'I also wanted to make sure everyone felt like part of the team......'

A careful distinction was made in the third step between Chinese numeral classifiers and measure words, as both numeral classifiers and measure words are tagged in BCC by 'q' (quantifiers). The distinction was made based on semantic criteria concerning *gè*-substitution, deletion in noun phrases, and cooccurrence with more than one noun referents, and one formal criterion relating to the distribution of adjectives in noun phrases, as discussed in §3.2.3. To be more specific, numeral classifiers can usually be replaced by *gè* or deleted in noun phrases without leading to changes in meanings of noun phrases, while the substitution of measure words with *gè* or their omission can lead to different interpretations of noun phrases, typically in terms of quantity. Furthermore, measure words, instead of numeral classifiers, are more likely to be used with more than one noun referent. Thirdly, the distribution of adjectives can influence the interpretation of noun phrases with measure words rather than numeral classifiers, as discussed in §3.2.3.

(62) Use of *bǎ* as a numeral classifier and a measure word

a. | *yī* | *bǎ* | *yǐzǐ* |
    |------|-------------|--------|
    | one | CLF:HANDLE | chair |

'one chair'

b. | *yī* | *bǎ* | *shāzǐ* |
    |------|---------------|---------|
    | one | MENS:HANDFUL | sand |

'one handful of sand'

Furthermore, polysemes and homophones were also examined and measure words were removed. As shown in example (62), *bǎ* in (a) is a numeral classifier and in (b) is a measure word, and therefore, the second one was removed from the results.

5) Two self-compiled corpora

Based on the data obtained from the above steps, two specialized corpora of numeral noun phrases were compiled. Corpus 1 was compiled based on the sequence of [*yī* 'one'+CLF+N], and consists of 6,700 pairs of Chinese-English numeral noun phrases without adjectives. In contrast, Corpus 2 was compiled based on the sequence of [*yī* 'one'+ADJ+CLF+N] and consists of 523 pairs of Chinese-English numeral noun phrases with adjectives.

4.2.2. Data analysis

Quantitative studies aimed to address research questions mainly concerning the interplay of different elements of numeral noun phrases, their semantic contributions, the semantic functions of numeral classifiers and their representation in English translation. In order to address the issues, the occurrence of numeral classifiers and other elements in Mandarin Chinese and their corresponding forms in English was calculated in the first step of data analysis. To be more specific, the frequency of different types of numeral classifiers was counted in the context of the absence and presence of adjectives, to show to what extent they are used in Chinese. Second, the proportions of the collocations of different types of numeral classifiers and different categories of nouns as well as adjectives were also calculated to examine their semantic contributions to noun phrases. Third, in order to show how properties specified or ascribed by numeral classifiers are reflected in English, the proportions of measure words in English were calculated as direct equivalents of numeral classifiers. For those numeral classifiers being omitted in English translation, different categories of nouns were compared to show whether additional properties expressed by numeral classifiers in Chinese were reflected by nouns in English. Fourth, the use of classifiers for the semantic function concerning individuation and the discourse function concerning definiteness was examined by calculating the proportion of the equivalents of the numeral *yī* 'one', together with the examination of singular and plural forms of head nouns in English, typically when numeral classifiers were omitted in translation. For example, by calculating the proportion of the numeral and pronoun *one* and the indefinite article *an* in English, the semantic unit created by numeral classifiers could

be reflected. By calculating the proportion of articles, on the other hand, one can show to what degree numeral classifiers are used to express definiteness.

4.3. Research methods for qualitative studies

Qualitative studies focused on the discourse functions of Chinese numeral classifiers and their representation in English translation. Based on an analysis of numeral classifiers in their collocates and contexts in discourse, the discourse functions of Chinese numeral classifiers were determined.

Qualitative studies were conducted based on the novel *The Three-Body Problem* by Liu Cixin ([2008] 2014). The novel was chosen as the data source for the following reasons. First, numeral classifiers occur more frequently and in a greater variety in fiction, according to Xiao and McEnery (2010: 49) and Erbaugh (1986: 403). Second, *The Three-Body Problem* is the first Asian science fiction novel ever to win the Hugo Award for Best Novel in 2015 and is also well-acknowledged as one of the best science fiction novels both in China and Western countries. Third, the novel can reflect the relatively recent use and translation of numeral classifiers in modern Chinese, as it was first published in 2008, and the English translation was published in 2014. Finally, both the writer and the translator of the novel have a Chinese background. The author Cixin Liu is a native Chinese, and the translator Ken Liu is a Chinese-American. They are both well-acknowledged writers in Chinese or English.

Selected chapters of the novel were chosen for analysis, i.e., chapters 1-5 of about 33,200 characters in the Chinese version and their equivalent chapters 4-8 of about 21,250 words in the English version.[16] These chapters were chosen for the study, as they serve as the introduction in the novel and are consistent in terms of narrative perspectives. These chapters were used to introduce the main characters, Wang Miao, Shi Qiang (nicknamed Da Shi), and Ye Wenjie, and other minor characters.

The data for qualitative studies were collected and cleaned in the following steps. First, the texts of the five chapters in Chinese and English versions were manually aligned

---

[16] The chapters in the English version were rearranged. The original chapters 1-5 taking place in 2005 in the Chinese version are shown as chapters 4-8 in Part II in the English version, and chapters 6-8 in the Chinese version are chapters 1-3 in Part I in the English version serving as a background taking place in 1967 during the Cultural Revolution.

before being imported into a parallel concordancing tool of BFSU ParaConc 1.2.1 (Xu et al. 2012). Second, the queries for noun phrases were made in BFSU ParaConc 1.2.1. In order to include all possible classifier phrases with and without numerals, nouns and their modifiers, the first round of queries were made based on the numeral classifiers derived from BCC for quantitative studies together with other possible numeral classifiers collected from studies of Chinese classifiers. Other numeral noun phrases without classifiers were also collected based on queries for numerals and demonstratives, in order to obtain a fuller picture of how noun referents are managed in discourse. Third, data cleaning was processed by focusing only on polysemous and homophonous morphemes. Numeral classifiers were distinguished from measure words or other particles. For example, similar to the cleaning step in §4.2.1, *bǎ* 'CLF:HANDLE' was distinguished from *bǎ* 'MENS:HANDFUL' as well as from *bǎ* as an object marker. Fourth, numeral noun phrases were also annotated. Aside from the five categories of numeral/demonstrative, adjective, numeral classifier, modifier, and noun, as in §4.2.1, noun referents were also identified based on context and annotated as a separate category. For example, one of the main characters, Wang Miao, is referred to as *rén* 'person', *xuézhě* 'academic', and *shèyǐng àihǎozhě* 'photography enthusiast' in four numeral noun phrases, and all these noun phrases were annotated with Wang Miao in terms of their referent. As a result, Corpus 3 was compiled, including 411 pairs of sentences and 645 pairs of noun phrases, consisting of at least two elements of numerals or demonstratives, numeral classifiers and nouns.

4.4. Concluding remarks

The chapter has given an account of the research methods applied in the dissertation, i.e., quantitative and qualitative studies. Quantitative studies were conducted based on 6,700 pairs of numeral noun phrases without adjectives and 523 pairs of numeral noun phrases in the context of adjectives from BCC. In comparison, qualitative studies were conducted based on 642 pairs of noun phrases in 411 pairs of sentences derived from the first five chapters of the novel *The Three-Body Problem*.

  While the three corpora were compiled for quantitative and qualitative studies, there were some limitations as to the source of data and types of numeral noun phrases in Corpus 1 and Corpus 2, the unbalanced size of the three corpora, and the variety of

numeral classifiers typically in Corpus 3. First, numeral noun phrases were compared between English and Chinese without indicating the source language of translation, for the online version of BCC, on which Corpus 1 and Corpus 2 were based, does not provide the source of its data or indicate the source languages of translation. Therefore, the collected data do not show in which direction numeral noun phrases in Corpus 1 and Corpus 2 are translated or how semantic and grammatical meanings of different elements in noun phrases are interpreted in the target language. Second, Chinese numeral noun phrases in Corpus 1 and 2 were strictly limited to two types of structures [*yī* 'one'+CLF+(MOD)+N] and [*yī* 'one'+ADJ+CLF+(MOD)+N]. As a result, they excluded numeral noun phrases with other numerals than *yī* 'one', demonstratives and interrogative pronouns, and classifier phrases not followed by nouns. Third, the size of the three corpora is unbalanced, as significantly fewer numeral classifiers were preceded by adjectives and fewer numeral noun phrases could be derived from the five chapters in *The Three-Body Problem*. Finally, not all numeral classifiers listed in the literature[17] were found in the three corpora due to their size. In spite of these limitations, however, these three corpora are sufficient to address current research questions about the semantics and functions of Chinese numeral classifiers in the following chapters.

---

[17] Lists of Chinese numeral classifiers were provided by e.g., Chao (1968: 598-609), Erbaugh (1986), Gao and Malt (2009: 1171-1177), Wu (2014b: 63-136), Ma (2015: 31), and Song (2017: 43-161).

Chapter 5: Chinese numeral classifiers: semantic contribution

5.1. Introduction

Much attention has been devoted to Chinese numeral classifiers concerning their features and functions, as shown in Chapter 3. However, relatively little systematic research has been conducted on the semantics and functions of Chinese numeral classifiers and their representation in English translation, typically based on corpus data. This chapter is devoted to addressing the first issue concerning the semantic contribution of different types of Chinese numeral classifiers and other elements of classifier phrases, based on the premise that different types of numeral classifiers are used in complementary distribution to express different meanings in relation to noun referents. To be more specific, this chapter aims to address the following research questions:

a)   To what extent are Chinese numeral classifiers used with and without adjectives in Chinese?

b)   How do Chinese numeral classifiers collocate with different types of nouns?

c)   How do Chinese numeral classifiers and other elements contribute to the semantics of the noun phrases?

To address the issues, I will first deal with the frequency of Chinese numeral classifiers in the two corpora of numeral noun phrases with and without adjectives in §5.2. In §5.3, collocations of numeral classifiers with different types of nouns are examined in Corpus 1. Based on the results, Chinese numeral classifiers will be examined based on the individuation hierarchy in this section. In §5.3.4, numeral classifiers will be examined with the collocation of nouns in the context of adjective in Corpus 2. The semantic

contribution of Chinese numeral classifiers and other elements of the classifier phrase will be discussed in §5.5 and §8.2.

5.2. Frequency of Chinese numeral classifiers in Corpus 1 and Corpus 2

While there is a large number of numeral classifiers in Chinese, they occur with different frequencies. Some numeral classifiers are used more frequently and collocate with a wider range of nouns, while others are used far less frequently and cooccur with only a limited scope of nouns. This section examines the frequency of numeral classifiers with and without adjectives in the two corpora of Chinese noun phrases.

The frequencies of all Chinese numeral classifiers have been calculated based on the occurrences of numeral classifiers as well as their types. As shown in §3.2.2, *gè* 'CLF:GENERAL' and *zhǒng* 'CLF:KIND, GENERAL' are regarded as a general classifier and a general kind classifier, while the other numeral classifiers are regarded as specific classifiers. Aside from the distinction between general and specific classifiers, there are also three other types of numeral classifiers: entity, event and kind classifiers. While *gè* 'CLF:GENERAL' can be regarded as an entity classifier and *zhǒng* 'CLF:KIND, GENERAL' can be regarded as a kind classifier, the two general classifiers are treated independently, since they occur significantly more frequently than any other numeral classifiers, which will be demonstrated below. Therefore, five types of Chinese numeral classifiers are here distinguished: the general classifier *gè*, the general kind classifier *zhǒng*, specific entity classifiers, specific event classifiers, and specific kind classifiers.

In Corpus 1 of 6700 noun phrases without adjectives, there are 104 Chinese numeral classifiers.[18] Their distribution in terms of the types mentioned above is shown in Table 10, in which the type/token difference is shown in the number of numeral classifiers within a given type and the number of their occurrences or tokens in the corpus. As can be seen, the two general classifiers are used far more frequently and account for almost 60% of all the occurrences of numeral classifiers. Specific classifiers constitute the remaining c. 40% of classifier tokens, with 82 specific classifiers (33.70%), 21 event classifiers (5.63%), and 3 kind classifiers (0.75%). The results indicate that general and

---

[18] For a complete list of Chinese numeral classifiers in Corpus 1, see 0

specific classifiers have a complementary distribution, with the two general classifiers used as default classifiers as they appear far more frequently, while specific classifiers are most likely to be used to express more specific meanings, considering their great variety.

Table 10. The frequency of different types of Chinese numeral classifiers in Corpus 1

type of numeral classifiers	no. of numeral classifiers	occurrence of numeral classifiers	
		#	%
general	2	4015	59.93
specific entity	82	2258	33.70
specific event	21	378	5.63
specific kind	3	50	0.75
total	104	6700	

The complementary distribution of Chinese numeral classifiers can be further attested in the semantics expressed by individual numeral classifiers. Take the top ten numeral classifiers in Table 11 as an example. As we can see, there is not much overlap in terms of the meanings they express. The most frequently used classifier, *gè,* is the least transparent in terms of its semantics, while *zhǒng* is a general classifier to express kind or type, two aspects that do not reflect salient features of noun referents.

Table 11. The top ten most frequent Chinese numeral classifiers in Corpus 1

numeral classifier	gloss	occurrence	
		#	%
gè	'general'	3111	46.43
zhǒng	'kind, general'	904	13.49
wèi	'individual, respect'	279	4.16
míng	'identity'	198	2.96
zhī	'single'	167	2.49
tiáo	'slender'	152	2.27
jiā	'household'	143	2.13
xiàng	'item'	133	1.99
cháng	'venue'	129	1.93
cì	'time'	100	1.49

In contrast, the other eight specific classifiers are more explicit with regard to the properties they denote. They are used to express properties of noun referents concerning humanness, animacy, and shape, as well as other meanings related to less frequently mentioned categories, e.g., organization, project, and venue and time of an event. Overlap may occur between *wèi* 'CLF:INDIVIDUAL, RESPECT' and *míng* 'CLF:IDENTITY', as they both denote humanness. However, they are different in affective meanings, with *wèi* 'CLF:INDIVIDUAL,

RESPECT' expressing more respect. What should also be noted is that the eight specific classifiers also include two specific event classifiers, although event classifiers constitute only a minor part of the system of Chinese numeral classifiers and in their occurrences in the corpus.

Compared with numeral classifiers used without adjectives, far fewer numeral classifiers are used in the context of adjectives in Corpus 2. As shown in Table 12, there are only 35 numeral classifiers, among which 34 are specific classifiers (over 99%).[19] The 29 specific entity classifiers constitute a predominant group, with about 96.75% of all occurrences. In contrast, the general classifiers and the other specific classifiers rarely occur with adjectives. For example, the general classifier *gè* occurs only four times with adjectives, while the general classifier *zhǒng* 'CLF:KIND, GENERAL' is not pre-modified by adjectives.

Table 12. The frequency of different types of Chinese numeral classifiers in Corpus 2

type of numeral classifiers	number of numeral classifiers	occurrence of numeral classifiers	
		#	%
general classifiers	1	4	0.76
specific entity	29	506	96.75
specific event	5	7	1.34
specific kind	1	6	1.15
Total	35	523	

Furthermore, among the 29 specific numeral classifiers, shape classifiers, i.e., specific numeral classifiers denoting shape including *kuài* 'CLF:LUMP.LIKE' (52.77%) and *piàn* 'CLF:FLAT/THIN' (17.21%), constitute the majority (92.73%). The results show that adjectives can restrict the choice of numeral classifiers. This can be attributed to the quantity reading of adjectives and shape classifiers. For example, the adjectives *dà* 'big, large' and *xiǎo* 'small, little' are modifiers denoting both size and quantity, and they are more likely to cooccur with numeral classifiers that are delimitable in terms of size, as with shape classifiers. This topic will be further discussed in §5.4.

In summary, Chinese general and specific classifiers occur in a complementary distribution when they are not modified by adjectives. General classifiers are used significantly more frequently, while a large number of specific classifiers are used to express more explicit features of noun referents. In the context of adjectives, however, shape classifiers are used significantly more frequently, which shows that the presence of adjectives

---

[19] For a list of Chinese numeral classifiers used with adjectives in Corpus 2, see Appendix 2.

affects the choice of numeral classifiers. The semantic contribution and distribution of numeral classifiers will be examined in collocations with nouns and adjectives in §5.3 and §5.4.

5.3. Collocations of Chinese numeral classifiers in Corpus 1

This section will examine Chinese numeral classifiers in collocation with different types of nouns and adjectives, with the aim to show their semantic contribution to noun phrases and their relationship in terms of the animacy hierarchy (Corbett 2000:55-56) and the individuation hierarchy (Audring 2006: 102; 2008: 107) (see §2.3.2.1). I will first discuss the collocations of numeral classifiers with nouns in Corpus 1 in §5.3, and then their collocations with adjectives and nouns in Corpus 2 in §5.4. The semantic contribution of numeral classifiers will be discussed in § 5.5.

5.3.1. Collocations of entity, kind and event classifiers

As described in §4.2.1, nouns were annotated in terms of four oppositions: human vs. nonhuman, animate vs. inanimate, concrete vs. abstract, and countable vs. uncountable. Table 13 shows the distribution of entity, kind and event classifiers with regard to these four oppositions in Corpus 1. Total 1 is the total number of different types of nouns. Total 2 is the total occurrence of different types of classifiers.

Table 13. The distribution of entity, kind and event classifiers in Corpus 1

semantic group	entity classifiers		kind classifiers		event classifiers		total 1
	#	%	#	%	#	%	
+human	1116	99.64	3	0.27	1	0.09	1120
-human	4209	76.04	950	17.16	376	6.79	5535
±human	44	97.78	1	2.22	0	0	45
+animate	1280	97.49	31	2.36	2	0.15	1313
-animate	4046	75.74	921	17.24	375	7.02	5342
±animate	43	95.56	2	4.44	0	0	45
+concrete	3584	90.53	309	7.78	66	1.67	3959
-concrete	1584	64.36	581	23.61	296	12.03	2461
±concrete	201	71.79	64	22.86	15	5.36	280
+count	4777	87.68	361	6.63	310	5.69	5448
-count	42	26.42	113	71.07	4	2.52	159
±count	550	50.32	480	43.92	63	5.76	1093
total 2	5369	80.13	954	14.24	377	5.63	6700

As shown in Table 13, nouns tend to collocate with entity classifiers (80.13%), including the general classifier *gè*, which may explain why entity classifiers are usually regarded as the only group of numeral classifiers in Chinese in most descriptions, as mentioned in §3.2.2.[20] As regards the distribution of entity classifiers, they cooccur predominantly with human and animate nouns (97.49%-99.64%). Slightly lower proportions are found among the collocations with concrete and count nouns (87.68%-90.53%) and with abstracts and the nouns of both concrete and abstract readings (64.36%-71.79%), e.g., *jiǎosè* 'character, role' and *biānjí* 'edit, edition, editor'. Only 26.42% of uncountable nouns cooccur with entity classifiers. However, the results challenge the argument that uncountable nouns cannot cooccur with numeral classifiers, as proposed by such scholars as Her (2012: 20). Uncountable nouns constitute a very small proportion of 2.34% but are shown in collocations with numeral classifiers, typically with kind classifiers, which will be discussed below.

Compared with entity classifiers, kind classifiers (14.24%), including the general kind classifier *zhǒng*, constitute a minor proportion of the collocations with nouns. As mentioned above, kind classifiers constitute most collocations with uncountable nouns (71.07%), and they also constitute 43.9% of the collocations with nouns annotated by [±count], e.g., *yánjiū* 'study, research' and *gōngzuò* 'job, work'.

The results show that kind classifiers are used in a complementary way with entity classifiers. Entity classifiers tend to cooccur with countable nouns (87.68%), while kind classifiers are more likely to be used with uncountable nouns (71.07%). Human and animate nouns are used almost exclusively with entity classifiers, with only 0.27%-2.36% of them collocated with kind classifiers. Among the remaining oppositions, kind classifiers are used a little more frequently with abstract, inanimate, and nonhuman nouns (17.24%-23.61%). Finally, only 6.63%-7.78% of kind classifiers are used with concrete and countable nouns, which are more likely to cooccur with entity classifiers. Further discussion about the complementary distribution of entity and kind classifiers will be given in §5.3.2 and §5.3.3.

Event classifiers make up the smallest group in the collocations with nouns. They tend to be used with abstracts (12.03%) instead of concrete nouns (1.67%). As regards the opposition of count vs. noncount, event classifiers are more likely to cooccur with

---

[20] See, e.g., Erbaugh (1986), Wu (2014b), Ma (2015), and Gao and Malt (2009).

countable nouns (5.69%) than uncountable nouns (2.52%). Finally, event classifiers are rarely collocated with human and animate nouns (0.09-0.15%), with only one or two collocations found in the corpus.

In summary, entity classifiers make up the largest share in collocations with nouns, typically countable concrete nouns. Kind classifiers complement entity classifiers by cooccurring more frequently with uncountable nouns. Event classifiers, used to denote features concerning events and processes, are more likely to be used with countable abstracts. The three types of Chinese numeral classifiers can thus be associated with three types of nouns: entity classifiers with countable concrete nouns, event classifiers with countable abstract nouns, and kind classifiers with uncountable nouns.

Based on these associations, the two general classifiers will be examined first in §5.3.2 in terms of their semantic features, especially concerning individuation. Specific classifiers will then be dealt with in §5.3.3 in their collocations with two types of countable nouns: concrete and abstract, as well as with nouns used as both countable and uncountable. In §5.3.4., the individuation hierarchy of Chinese numeral classifiers will be examined based on the findings in §5.3.2 and §5.3.3. Finally, §5.3.5 will be devoted to the variation in the use of Chinese numeral classifiers, where, e.g., entity classifiers are used with uncountable nouns and event classifiers are used with concrete nouns.

5.3.2. Collocations of general classifiers and specific kind classifiers

This section turns to the collocations of the two general classifiers in Corpus 1: *gè* 'CLF:GENERAL' and *zhǒng* 'CLF:KIND, GENERAL'. As specific kind classifiers constitute a very small part in the collocations with nouns, their collocations will also be examined in this section.

Table 14. The distribution of the two general classifiers and specific kind classifiers in Corpus 1

semantic group	*gè*		*zhǒng*		specific kind classifiers		total 1
	#	%	#	%	#	%	
+animate	663	50.50	28	2.13	3	0.23	1313
-animate	2418	45.26	874	16.36	47	0.88	5342
±animate	30	66.67	2	4.44	0	0	45
+human	649	57.95	1	0.09	2	0.18	1120
-human	2432	43.94	902	16.30	48	0.87	5535
±human	30	66.67	1	2.22	0	0	45
+concrete	1602	40.46	282	7.12	27	0.68	3959
-concrete	1343	54.57	566	23.00	15	0.61	2461
±concrete	166	59.29	56	20.00	8	2.86	280
+count	2730	50.09	323	5.93	38	0.70	5449
-count	4	2.55	106	67.52	7	4.43	158
±count	377	34.49	475	43.46	5	0.46	1093
total 2	3111	46.43	904	13.49	50	0.75	6700

Table 14 shows the distribution of the two general classifiers and specific kind classifiers in terms of the four semantic oppositions. In the table, total 1 is the total number of all numeral classifiers used with relevant types of nouns and total 2 is the total number of the occurrence of one type of numeral classifiers. The general classifier *gè* is distributed relatively evenly among collocations with different groups of nouns. It is more frequently used with nouns with unspecified features and annotated as ±animate and ±human (66.67%). However, the two groups of nouns are quite limited in number (45). Regarding other types of nouns, *gè* is more likely to cooccur with countable, abstract, and animate nouns (50.09%-54.57%) instead of concrete, nonhuman and inanimate nouns (40.46%-45.26%). However, a closer look into the collocations of *gè* with animate nouns also shows that *gè* is almost exclusively used with nouns for humans. Among the 663 cooccurrences with animates, 649 of them are used with human nouns. Therefore, *gè* is a general classifier for countable nouns. However, when used with animate nouns, *gè* tends to denote humanness.

Only 4 in 3111 of the occurrences of *gè* are found in the collocations with uncountable nouns, a significant lower percentage of 2.55%. As *gè* denotes a discrete unit of nouns, it is rarely used with uncountable nouns with unspecified features in terms of individuation. Uncountable nouns (67.52%) are more likely to cooccur with the general classifier *zhǒng* 'CLF:KIND, GENERAL'. Furthermore, *zhǒng* can also replace the general classifier *gè* and other kind classifiers when used with uncountable nouns. For example,

while *kàngyǎnghuàjì* 'antioxidant' appears with both the general classifier *gè* and the general kind classifier *zhǒng*, the latter is regarded as a more natural choice.

The classifier *zhǒng* can also be used with countable nouns (5.93%), as shown in Table 15. In this context, it is more likely to cooccur with countable nouns annotated as ±concrete (19.03%) than countable abstract nouns (7.68%) and countable concrete nouns (4.38%).

Table 15. The distribution of the general kind classifier *zhǒng* among countable nouns

countable nouns	*zhǒng*		total 1
	#	%	
+concrete	164	4.38	1458
-concrete	112	7.68	2461
±concrete	47	19.03	247
total 2	323	5.93	5448

When collocated with countable nouns annotated by ±concrete and abstract nouns, the general kind classifier *zhǒng* can usually be replaced by the general classifier *gè*, as with *zhuāngzhì* 'installation, device, equipment', *xìtǒng* 'system' and *tújìng* 'way'. Referents of these nouns tend to be unobservable either in space or time. However, they are not regarded as uncountable nouns that can be modified by such adverbs as *yīdiǎndiǎn* 'a little bit'. Countable concrete nouns collocated with the general kind classifier *zhǒng* are usually such hypernyms as *zhíwù* 'plant', *bìngdú* 'virus', and *gōngjù* 'tool'. Therefore, it expresses a collective meaning of the referents, which confirms that kind classifiers can coerce a kind or type reading of noun referents (Huang and Ahrens 2003: 17-23).

Three specific kind classifiers constitute the minimal share (0.75%) of the collocations with nouns, as shown in Table 14. They are more likely to cooccur with countable nouns. Among the 50 tokens of specific kind classifies, 38 of them appear with countable nouns. They express more specific features of a type of referents. For example, *kuǎn* 'CLF:STYLE' tends to refer to a type of products, as with *chéngxù* 'program', *qiǎokèlì* 'chocolate', and *chǎnpǐn* 'product'. In this context, they cannot be replaced by the general classifier *gè*, and the replacement with *zhǒng* may be less specific as to their referents.

To conclude, *gè* is used as a default classifier for countable nouns, as discussed in §3.2.2, as it denotes a discrete unit of countable nouns without showing specific features of their referents. In contrast, uncountable nouns, which are less discrete in terms of

individuation, are more frequently collocated with *zhǒng* or measure words[21]. When used with countable nouns, *zhǒng* expresses a type or kind instead of a unit. Among the collocations with animate nouns, *gè* tends to express humanness. In contrast, *zhǒng* is rarely collocated with specific animate nouns for either animals or humans.

5.3.3. Collocations of specific classifiers

This section focuses on the collocations of specific classifiers with nouns, involving the four semantic oppositions. The collocations of specific classifiers with human and animate nouns will be discussed first, followed by other countable nouns and nouns with both countable and uncountable readings.

5.3.3.1. Collocations with human nouns

Human nouns are predominantly used with entity classifiers (99.64%), with the remaining minimal part shared by event classifiers (0.09%) and kind classifiers (0.27%), as shown in Table 13. While event and kind classifiers can be used with human nouns, they do not indicate inherent properties of human referents. Instead, they ascribe additional meanings to the referents, which will be further discussed in §5.3.5. This section focuses on the collocation of entity classifiers with human nouns.

Table 16. The distribution of entity classifiers among human nouns[22]

entity classifier	gloss	occurrence	
		#	%
gè	general	679	58.64
weì	'individual'	279	24.09
míng	'identity'	198	17.10
dài	'generation'	2	0.17
total		1158	

---

[21] Previous studies show that uncountable nouns can only cooccur with measure words, see, e.g., Her (2012: 20).

[22] The data is derived based on human nouns annotated as [+human], and those distinguished from nouns annotated as [±human].

As shown in Table 16, other than the general classifier *gè*, there are three entity classifiers used to denote humanness: *wèi* 'CLF:INDIVIDUAL, RESPECT', *míng* 'CLF:IDENTITY', and *dài* 'CLF:GENERATION'. The general classifier *gè* accounts for more than a half (58.64%) of the collocations with nouns for humans, while *wèi* and *míng* constitute 24.09% and 17.10%, respectively, and *dài* makes up 0.17%. While *wèi* and *míng* are specialized for human nouns, *dài* 'CLF:GENERATION' is far less frequently collocated with human nouns. It is usually used with nouns whose referents are very famous or successful in a certain period of time, as with the noun *míngxiàng* 'famous.prime.minister'. Furthermore, it can also be used with nonhuman nouns, such as *zázhǒng* 'hybrid'.

The results confirm that *gè* is the general classifier for human nouns. First, it cooccurs far more frequently with human nouns, which means it is more likely to be used as a default classifier with human nouns. Second, it cooccurs with a broader range of human nouns, including general ones, including *rén* 'person' and *jiāhuó* 'thing, guy', and more specific ones, including nouns for humans of a relatively higher social status, e.g., *guówáng* 'king', *guānyuán* 'official', *jiàoshī* 'teacher', *jǐngchá* 'policeman', and *fùqīn* 'father', as well as those of a relatively lower social status, e.g., *qǐgài* 'beggar', *yuángōng* 'employee', *xuéshēng* 'student', *qiángdào* 'robber', and *háizǐ* 'child'. However, if nouns for humans of higher social status are used with the general classifier *gè*, it may imply a derogatory meaning, typically when the referent in question is present. Furthermore, it can also be used with kinship terms and titles of address, e.g., *māmā* 'mom', *biǎomèi* 'cousin', and *lǎoshī* 'teacher'. The variety of nouns collocated with *gè* confirms the prediction about the unspecified semantics of *gè* discussed in §3.2.2 above.

While *wèi* 'CLF:INDIVIDUAL, RESPECT' can be used with most human nouns, similar to *gè*, it is less likely to cooccur with generic nouns, informal nouns or nouns that imply a derogatory meaning. For example, when referring to the elderly, *wèi* 'CLF:INDIVIDUAL, RESPECT' is a more proper choice, as in collocations with *lǎorén* 'elder', *lǎowēng* 'greybeard', and *lǎotàitài* 'old lady'. However, it is not proper for such informal nouns as *lǎotóu* 'old man, codger' or such derogatory terms as *lǎotàipó* 'old woman'. Therefore, *wèi* 'CLF:INDIVIDUAL, RESPECT' is correlated with noun referents of higher social status, as discussed in §3.2.2, regarding the due respect it expresses to them. When *wèi* is used with nouns conveying a lower social status, e.g., *qǐgài* 'beggar', *xiǎohái* 'child', and *zuìfù* 'sinful woman', it ascribes respect to noun referents.

Compared with *gè* and *wèi, míng* 'CLF:IDENTITY' is a more specific human classifier used with nouns denoting occupations and social roles. For example, it is more likely to be used with *jiàoshī* 'teacher', which refers to the occupation, instead of *lǎoshī* 'teacher', usually used as a title of address, or *rén* 'person' with general identity. In Corpus 1, *míng* 'CLF:IDENTITY' occurs in five of the eight phrases with *jiàoshī* 'teacher'. However, it appears in none of the ten phrases of *lǎoshī* 'teacher' nor in the 64 phrases of *rén* 'person'. Furthermore, it is more likely to be used with such nouns for females as *fùnǚ* 'woman' and *nǚzǐ* 'female, woman' than with such kinship terms as *māmā* 'mom' and *nǚer* 'daughter'.

In summary, other than the general classifier *gè*, there are two main specific human classifiers: *wèi* 'CLF:INDIVIDUAL, RESPECT' and *míng* 'CLF:IDENTITY'. They do not differentiate genders among human nouns, as in gender systems in Afro-Asiatic and Indo-European languages, as discussed in §2.3. However, they are used to denote humanness typically with regard to *wèi* 'CLF:INDIVIDUAL, RESEPCT' and *míng* 'CLF:IDENTITY'. Furthermore, the two specific classifiers denote specific semantic features, such as occupations and social roles by *míng* 'CLF:IDENTITY' and affective meanings by *wèi* 'CLF:INDIVIDUAL, RESPECT'. Finally, the use of *wèi* 'CLF:INDIVIDUAL, RESPECT' and *gè* 'CLF:GENERAL' is related to social status, with *wèi* expressing more respect and *gè* possibly implying more contempt to referents.

5.3.3.2. Collocations with non-human animate nouns

While *wèi* 'CLF:INDIVIDUAL, RESPECT' and *míng* 'CLF:IDENTITY' are used almost exclusively for humans, other specific entity classifiers can be applied to other concrete nouns, such as nouns for animals, objects, and instruments. Table 17 shows the distribution of numeral classifiers among animal nouns.

Table 17. The distribution of numeral classifiers among animal nouns[23]

type of numeral classifiers	numeral classifier	occurrence	
		#	%
entity classifiers	*gè*	13	6.78
	specific	150	78.13
kind classifiers	*zhǒng*	27	14.06
	specific	1	0.52
event classifiers		1	0.52
total		192	

As indicated in the table animal nouns are most likely used with entity classifiers (78.13%), excluding the general classifier *gè* (6.78%). Among the remaining animal nouns, they cooccur more frequently with the general kind classifier *zhǒng* (14.06%) and are least likely to be used with event classifiers and specific kind classifiers, with only 1 cooccurrence out of 192 in total number of the collocations, respectively. The following discussion focuses on the collocation of specific entity classifiers with animal nouns, while their collocations with other numeral classifiers will be dealt with in §5.3.5.

Table 18. The collocation of entity classifiers with animal nouns

specific entity classifier	gloss	occurrence	
		#	%
zhī	'single'	111	73.51
tiáo	'slender'	20	13.25
tóu	'head'	9	5.96
pǐ	'horse, horse-like animal'	6	3.97
wěi	'tail'	2	1.32
jià	'framework'	1	0.66
jù	'long and stiff'	1	0.66
dài	'generation'	1	0.66
total		151	

As shown in Table 18 among eight entity classifiers used for animal nouns, *zhī* 'CLF:SINGLE' (73.51%) is used as the most frequent one. While *zhī* 'CLF:SINGLE' can also be used with nouns for objects, which will be discussed in §5.3.3.3, it is a more typical choice for animal nouns. *Zhī* is collocated widely with nouns for animals of various kinds, including mammals, e.g., *lǎohǔ* 'tiger', *hóuzǐ* 'monkey', and *gǒu* 'dog', reptiles, e.g., *kǒnglóng* 'dinosaur' and *shé* 'snake', amphibians, e.g., *qīngwā* 'frog' and *chánchú* 'toad', invertebrates, e.g., *hǎixīng* 'sea star', *xiè* 'crab', insects, e.g., *cāngyíng* 'fly', fish, e.g., *shāyú* 'shark', and birds of various kinds. The high frequency of *zhī* 'CLF:SINGLE' in collocation

---

[23] The data is derived based on animal nouns annotated as [+animate] [-human], and those distinguished from nouns annotated as [±animate] [±human].

with nouns for animals and its unspecified features confirm the prediction that *zhī* is the general classifier for nouns for animals, as mentioned in §3.2.2.

Other specific entity classifiers denote more specific features of nouns for animals. For example, *tiáo* 'CLF:SLENDER' can be used with such nouns as *yú* 'fish', *shé* 'snake' and *gǒu* 'dog', whose referents can also be described as long and slender, while *tóu* 'CLF:HEAD' can be used with nouns for animals whose heads show salient features, as in *lù* 'deer', *dàxiàng* 'elephant', and *niú* 'ox'.

Some numeral classifiers can only be used with nouns for specific animals or specific kinds of animals. For example, *pǐ* 'CLF:HORSE, HORSE-LIKE.ANIMAL' is only used with nouns for horses or horse-like animals, e.g., *mǎ* 'horse' and *luó* 'mule', and *wěi* 'CLF:TAIL' is typically used with nouns for fish. Other numeral classifiers denoting shape and dimension can cooccur with a wider range of nouns not only for animals but also for other concrete or abstract entities. For example, *zhī* 'CLF:SINGLE' and *tiáo* 'CLF:SLENDER' can also be used with inanimate nouns, such as *zhī* with *shǒu* 'hand' and *píngzi* 'bottle', and *tiáo* with *máojīn* 'towel', *gōnglù* 'highway', and *xìnxī* 'message' (see also §5.3.3.3 and §5.3.3.4 below).

To sum up, there are eight specific entity classifiers used with nouns for animals. The general classifier for animal nouns is *zhī* 'CLF:SINGLE'. There are no cases of the collocation of animal classifiers with human nouns in the corpus. Likewise, some animal classifiers, e.g., *pǐ* 'CLF:HORSE, HORSE-LIKE.ANIMAL' and *wěi* 'CLF:TAIL', are not applied to inanimate nouns. However, most animal classifiers, typically denoting shape and dimension, can also cooccur with other inanimates.

5.3.3.3. Collocations with countable concrete inanimates

Most entity classifiers occur with countable concrete nouns, typically inanimates. As shown in Table 19, the numeral classifiers most frequently collocated with nouns annotated by [+count] [+concrete] [-animate] are entity classifiers. While the general classifier *gè* is used as the default classifier for countable concrete inanimates (37.38%), nouns in this group are more likely to be differentiated based on features denoted by specific entity classifiers, as most specific entity classifiers (67 out of 82) are found to be collocated with this group of nouns and constitute 54.19% in the collocations.

Table 19. The distribution of numeral classifiers among countable concrete inanimates

type of numeral classifiers	numeral classifier	occurrence	
		#	%
entity classifiers	*gè*	900	37.38
	specific	1305	54.19
kind classifiers	*zhǒng*	136	5.65
	specific	16	0.66
event classifiers		51	2.12
total		2408	

The entity classifier *zhī* 'CLF:SINGLE', which is a general classifier for animal nouns , as shown in §5.3.3.2, is also used with some nouns in this group. Its unspecified feature allows it to be applied to various kinds of nouns, e.g., *jièzhǐ* 'ring', *shǒu* 'hand', *yǎnjīng* 'eye', *píngzǐ* 'bottle', *dòng* 'hole', *xié* 'shoe', and *xiāngjiāo* 'banana'. However, it constitutes a small proportion (4.29%) of the collocations with this group of nouns, and therefore, it cannot be regarded as a general classifier for countable concrete inanimates. Other less specified entity classifiers include *jiàn* 'CLF:PIECE', *fèn* 'CLF:SHARE, PORTION', and *zé* 'CLF:CLAUSE, ENTRY'. They can only be used with countable inanimates, including countable abstracts. Their collocation with countable abstracts will be discussed in §5.3.3.4.

Entity classifiers denoting shape, size and dimensionality are usually collocated with more varied types of nouns. For example, one of the animal classifiers *tiáo* 'CLF:SLENDER' is found to be widely assigned to concrete nouns featuring a long and slender shape, including nouns for roads, e.g., *mǎlù* 'road', *suìdào* 'tunnel', and *tiělù* 'railway', garments, e.g., *kùzi* 'trousers', accessories, *lǐngdài* 'tie', *shǒuliàn* 'bracelet', and *wéijīn* 'scarf', body parts, e.g., *tuǐ* 'leg' and *shǒubì* 'arm', part for plants, e.g., *néngzhī* 'branch, twig', and even equipment, e.g., *liúshuǐxiàn* 'assembly.line'. Similar shape classifiers include *zhāng* 'CLF:SPREADING.OPEN/FLAT', *kuài* 'CLF:LUMP.LIKE' and *kē* 'CLF:ROUNDISH'.

Entity classifiers denoting salient features, however, are more likely to be used with a limited range of nouns. For example, *jiā* 'CLF:HOUSEHOLD', which is one of the ten top specific classifiers, as listed in Table 11, is usually used with nouns for organizations and institutions, including factories, stores, restaurants, companies, clubs, banks, courts, and schools. Similar classifiers also include *běn* 'CLF:BOOK' for books, *piān* 'CLF:ARTICLE' for articles, *liàng* 'CLF:VEHICLE, CAR' for vehicles, *kē* 'CLF:PLANT' and *zhū* 'CLF:STALK' for plants, *dòng* 'CLF:BUILDING, BEAM' and *zhuàng* 'CLF:BUILDING, PILLAR' for buildings, and *bǎ* 'CLF:HANDLE' for instruments and household items which feature a handle.

In summary, entity classifiers are most likely collocated with nouns annotated as [+count][+countable][-animate]. While *gè* is the general classifier for this group of nouns, specific entity classifiers are more likely to be used to denote specific features concerning shape, dimensionality and other salient physical features, as discussed in §3.2.2. Shape classifiers tend to be more flexible and applied to more varied types of nouns within this group, while numeral classifiers denoting salient features are more likely to be used with one or two fixed categories of nouns. Specific entity classifiers usually do not cooccur with nouns for humans and animals. However, they may cooccur with abstract nouns, which will be discussed in §5.3.3.4.

5.3.3.4. Collocations with countable abstracts

There are two types of countable abstract nouns: 1) nouns for observable events, changes and processes; 2) nouns for non-observable concepts, as shown in §4.2.1. While the first group of nouns are observable in terms of duration of time or place, some nouns in the second group are also conceptually bounded. For example, languages can be distinguished from one another in speech or writing, while such concepts of time as *bāitiān* 'day, daytime' can be regarded as bounded based on the sunrise and the sunset. Table 20 shows the distribution of entity, kind and event classifiers among countable abstract nouns.

Table 20. The collocation of numeral classifiers with countable abstracts

type of numeral classifiers	numeral classifier	noun for non-observable concept		noun for observable event	
		#	%	#	%
entity classifiers	*gè*	916	80.21	84	26.92
	specific	75	6.57	14	4.49
kind classifiers	*zhǒng*	105	9.19	4	1.28
	specific	14	1.23	0	0
event classifiers		32	2.80	210	67.31
total		1142		312	

Among countable abstracts, there are more nouns for non-observable concepts than for observable events (1142 to 312). Entity classifiers are far more frequently used with nouns for non-observable concept, compared with the other two types of numeral classifiers. However, the general classifier *gè* constitutes the largest share of 80.21%, while

only 6.57% of specific entity classifiers are used with this group of nouns. A similar trend is also found in the use of general (9.19%) and specific kind classifiers (1.23%). As discussed in §5.3.2, kind classifiers, typically *zhǒng* 'CLF:KIND, GENERAL', can be replaced by the general classifier *gè*. These results show that nouns for non-observable concepts are far less specified, compared with those countable concretes, including for humans and animals, as discussed in § 5.3.3.1 to §5.3.3.3.

Event classifiers (210 out of 231), in contrast, cooccur far more frequently with nouns for observable events, while the general classifier *gè* constitute only 26.92% in the collocations and specific entity classifiers and kind classifiers are rarely used with this group of nouns (14 and 4 respectively). Among the 16 event classifiers, *cháng* 'CLF:VENUE' (28.9%) and *cì* 'CLF:TIME' (17.3%) are used far more frequently with a broader range of nouns for events. Furthermore, they are usually interchangeable with each other, as in collocations with *bǐsài* 'match'. Other event classifiers are used with a more specific group of nouns. For example, *jiàn* 'CLF:PIECE', *zhuāng* 'CLF:STAKE', and *zōng* 'CLF:FACTION' are more typically used with such nouns as *shì* 'thing', *gōng'àn* 'case' and *bìnggòu'an* 'deal', while *dùn* 'CLF:SPELL, SESSION', *táng* 'CLF:CLASS', and *bǐ* 'CLF:PEN' are more likely to be used with nouns for meals, classes and transactions, respectively. What should be noted is that *jiàn* 'CLF:PIECE' as a less specified entity classifier is more frequently used with such event nouns as *shì* 'thing' and *gōng'àn* 'case'.

The results suggest that the general classifier *gè* can still be regarded as a default classifier for countable abstracts. However, nouns for non-observable concepts are comparatively less likely to be specified by specific classifiers, while nouns for observable events tend to be specified by event classifiers instead of specific entity classifiers. Kind classifiers are not very typical for this group of nouns in general, and they tend to be replaceable with the general classifier *gè*.

5.3.3.5. Collocations with nouns annotated by [count] [±concrete]

Regarding countable nouns with both concrete and abstract referents, the general classifiers *gè* (61.54%) and *zhǒng* (19.03%) constitute a significant share of the collocations, as shown in Table 21, while specific classifiers make up the remaining part of less than 20%.

Table 21. The collocation of numeral classifiers with nouns annotated by [count] [±concrete]

type of numeral classifiers	numeral classifier	occurrence	
		#	%
entity classifiers	*gè*	152	61.54
	specific	28	11.34
kind classifiers	*zhǒng*	47	19.03
	specific	5	2.02
event classifiers		15	6.07
total		247	

Kind classifiers used with this group of nouns can also be replaced by *gè* without leading to much difference in the interpretation, as with *zhuāngzhì* 'device, instalment', *jiégòu* 'structure, construction', and *móxíng* 'model'. While specific classifiers can be used to distinguish whether the nouns are concrete or abstract, specific entity classifiers tend to specify concrete referents. As shown in example (63) [24], while *zhōngxīn* 'centre' can refer to a concrete organization or an abstract concept, the entity classifier *jiā* 'CLF:HOUSEHOLD' specifies that the referent is an organization.

(63) The use of specific entity classifiers with nouns of both concrete and abstract referents

yī	*jiā*	*zhōngxīn*
one	CLF:HOUSEHOLD	centre

'one centre'

In contrast, event classifiers tend to refer to events. For example, *tàocān* 'set meal, package' can refer to a combination of different food, a set meal, or a package, e.g., for internet data or a tour. In example (64), its referent is specified as a set meal due to the presence of the event classifier *dùn* 'CLF:SPELL, SESSION'.

(64) The use of event classifiers with nouns of both concrete and abstract referents

yī	*dùn*	*tàocān*
one	CLF:SPELL, SESSION	combo, set.meal, package

'one set meal'

To conclude, nouns annotated by [+count] [±concrete] tend to be less specified, and therefore, they are more likely used with the general classifiers, and *gè* can also replace the general kind classifier *zhǒng* in this context. Relatively smaller proportions of specific

---

[24] Examples and their English translations in this chapter are collected from Corpus 1 and Corpus 2, unless otherwise indicated.

entities and event classifiers are used with this group of nouns. Specific entity classifiers are usually used to refer to concrete entities, while event classifiers tend to be used to refer to events.

5.3.3.6. Collocations with nouns annotated by [±count]

The general classifier *gè* is more frequently used with countable nouns, as shown above, and the general kind classifier *zhǒng* is more typical for uncountable nouns, as shown in §5.3.2. However, for polysemous nouns that have both countable and uncountable readings, not much difference can be found in the use of the two general classifiers. As shown in Table 22, *gè* (34.49%) and *zhǒng* (43.46%) account for a significant proportion of the collocations with nouns annotated by [±count]. What should be noted is that over 90% of nouns in this group are abstract. Furthermore, *gè* and *zhǒng* tend to be interchangeable when used with this group of nouns, as with *shēngyīn* 'sound, voice', *fāngfǎ* 'way, method', and *dòngzuò* 'movement, action'.

Table 22. The collocation of numeral classifiers with nouns annotated by [±count]

type of numeral classifiers	numeral classifier	occurrence	
		#	%
entity classifiers	*gè*	377	34.49
	specific	173	15.83
kind classifiers	*zhǒng*	475	43.46
	specific	5	0.46
event classifiers		63	5.76
total		1093	

Furthermore, relatively more specific entity classifiers (15.83%) than specific event classifiers (5.76%) are used with this group of nouns. Considering their low proportions in the distribution, neither of them can be regarded as specialized for this group of nouns. Some of them are more general. For example, *xiàng* 'CLF:ITEM, PROJECT' (7.96%) is widely used with abstract nouns unspecified in terms of countability, e.g., *gōngzuò* 'job, work', *yánjīu* 'study, research', *àihào* 'hobby, enthusiasm', *quánlì* 'power, right', *chéng*guǒ 'product, achievement', and *jìhuà* 'plan, planning'. Others can be more specific to denote relevant features of entities and events.

(65) The use of specific classifiers with generic nouns
    a.   *yī*   *tiáo*   *zŭzhī*
        one   CLF:SLENDER   tissue, organization
      'one band of tissue'
    b.   *yī*   *cháng*   *jìhuà*
        one   CLF:VENUE   planning, plan
      'one activity of planning, one campaign'

The examples in (65) illustrate how specific classifiers are used to differentiate referents of nouns annotated as [±count]. *Zŭzhī* 'tissue, organization' can refer to body tissue or an organization. By being collocated with *tiáo* 'CLF:SLENDER' in (a), it refers to the former. Similarly, *jìhuà* 'plan, planning' refers to an activity of planning when used with *cháng* 'CLF:VENUE' in (b), although it generally refers to a plan.

    Therefore, nouns unmarked for countability are more likely to be used with the general classifiers *gè* and *zhŏng*. On the other hand, most specific classifiers are applied to this group of nouns based on the properties of their referents.

    This section has dealt with the collocations of specific classifiers, which have been comparisons with general classifiers. Specific entity classifiers are most frequently used with countable concrete nouns to denote humanness, animacy, physical and other salient features, while event classifiers cooccur predominantly with countable abstract nouns for observable events. Among specific classifiers, those denoting more salient physical features tend to be more fixed in the collocations with one or two groups of nouns. In contrast, shape classifiers and classifiers less specific tend to cooccur with a wider range of nouns for animals, concretes, and abstracts. Furthermore, specific classifiers can also be used to specify referents of general nouns annotated by [±count] and [±concrete] based on their salient properties. The results also show that while *gè* is a general classifier for most countable nouns, *zhī* should be regarded as a default classifier for animal nouns. Finally, the less specific nouns are, the more likely that *gè* and *zhŏng* are used interchangeably with them.

5.3.4. The individuation hierarchy

While the results of the above corpus study have shown a semantic correlation of Chinese numeral classifiers and their head nouns, they also indicate that their distribution is

closely associated with the parameter of individuation. As mentioned in §2.3.2.1, this parameter is related to the countability of noun referents in terms of graded degrees (Fletcher 1987; Audring 2006: 94) (see Fig. 13).

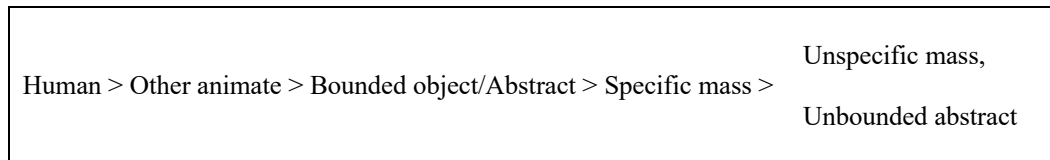|	Unspecific mass,
Human > Other animate > Bounded object/Abstract > Specific mass >	Unbounded abstract

Fig. 13. Individuation Hierarchy (Audring 2006: 102).

In this hierarchy, types of noun referents are ordered in accordance with their degree of individuation. Humans and animals are ordered high in the hierarchy, as they can be clearly identified. Most concrete inanimates are distinguishable in terms of physical boundedness, and some abstract concepts are directly observable in terms of duration of time or are conceptually bounded. Therefore, these two groups of referents are ordered in the middle in the hierarchy. Lowest in the hierarchy are entities that are fuzzy in their physical boundedness or concepts unobservable in either time or space.

The distribution of entity and kind classifiers, typically the two general classifiers *gè* 'CLF:GENERAL' and *zhǒng* 'CLF:GENERAL, KIND', is closely related to the countability of nouns (see Table 13 and Table 14 above). The classifier *gè* occurs predominantly with countable nouns (87.68%), typically with nouns for humans (58.64%) and nouns for bounded concepts (80.21%), while *zhǒng* tends to occur in collocations with uncountable nouns (71.07%). In collocations with nouns marked as both countable and uncountable, *gè* (34.49%) and *zhǒng* (43.36%) occur less frequently and they tend to be interchangeable (see Table 26).

Specific classifiers tend to be used with different types of countable nouns based on their semantics. For example, specific human classifiers are used only with human nouns, and animal classifiers based on salient physical features cooccur only with nouns for animals. Among human nouns, *wei* 'CLF:INDIVIDUAL, RESPECT' and *gè* can show a different status of their referents. Nouns for animals are differentiated from human nouns by being more frequently collocated with *zhī* 'CLF:SINGLE' instead of *gè*.

Most specific entity classifiers cooccur with nouns for bounded entities (54.19%), other than nouns for humans and animals, as shown in §5.3.3.3. Among specific entity classifiers for bounded entities, most classifiers based on salient physical features can

only be used with specific categories of nouns, e.g., *jiā* 'CLF:HOUSEHOLD' for organizations and institutions, *kē* 'CLF:PLANT' and *zhū* 'CLF:STALK' for plants, *dòng* 'CLF:BUILDING, BEAM' and *zhuàng* 'CLF:BUILDING, PILLAR' for buildings, and *běn* 'CLF:BOOK' for books, as mentioned in §5.3.3. Similarly, most event classifiers are used with nouns for bounded events, or observable events (67.31%), as shown in §5.3.3.4. However, specific entity classifiers are less frequently used with nouns for non-observable concepts (10.60%), as their features are less salient. Instead, these nouns are predominantly used with the general classifier *gè* (80.21%), as shown in Table 20. For nouns for entities unspecified in terms of countability or concreteness, specific classifiers constitute only 19.43% to 22.05% in all collocations, as shown in Table 24 and Table 26. Specific classifiers constitute 29.93%, still a minor proportion, in the collocations with uncountable nouns (see also §5.3.5 below).

As Chinese numeral classifiers are used to individuate nouns, as discussed in §3.3.1, and they are also semantically correlated with noun referents, as shown in §3.2.2 and §5.3, they can also show the graded degrees of individuation on the hierarchy. Table 23 shows the distribution of the different types of numeral classifiers in terms of the individuation hierarchy.

Table 23. Individuation hierarchy and semantic correlation of Chinese numeral classifiers

individuation hierarchy		general classifiers	specific classifiers
Human		*gè*	entity classifiers for humans
Animal		*zhī*	entity classifiers for animals and shape classifiers
Bounded inanimate	entity	*gè*	entity classifiers
	event	*gè*	event classifiers
	concept	*gè*	less specified entity classifiers and shape classifiers
Neutralized entity		*gè, zhǒng*	entity, event, and kind classifiers
Mass		*zhǒng*	kind classifiers

The general classifier *gè* and specific human classifiers are ordered at the top of the hierarchy for human referents, followed by the general classifier *zhī* 'CLF:SINGLE' and specific entity classifiers for animal referents. Lowest in the hierarchy is the general kind classifier *zhǒng* for mass nouns. Most specific entity classifiers, together with the general classifier *gè*, are ranked in the middle of the hierarchy and assigned to nouns of different referents based on their semantics. For example, entity classifiers based on salient features are used for bounded entities, while event classifiers are used for bounded events.

What should also be noted is that more general entity classifiers and shape classifiers are more flexible by being assigned to more varied types of nouns. For example, less specified entity classifiers, e.g., *jiàn* 'CLF:PIECE' and *fèn* 'CLF:SHARE, PORTION', can be used with both concrete and abstract inanimates. As shown in example (66), *jiàn* 'CLF:PIECE' can be used with concrete nouns, as with the more general one *dōngxi* 'thing' in (a) and the more specific one *chènyī* 'blouse, shirt', and it can also cooccur with abstract nouns, as with the more general one *shì* 'thing' in (c) and the more specific one *chéngjìu* 'success, achievement' in (d).

(66) Collocations of *jiàn* with concrete and abstract inanimates

a. | *yī* | *jiàn* | *tèbié* | *de* | *dōngxi* |
    | one | CLF:PIECE | special | MOD | thing |
    'something special'

b. | *yī* | *jiàn* | *báisè* | *chènyī* |
    | one | CLF:PIECE | white | blouse, shirt |
    'a white blouse'

c. | *yī* | *jiàn* | *shì* |
    | one | CLF:PIECE | thing |
    'one thing'

d. | *yī* | *jiàn* | *xiāngdāngkěguān* | *de* | *chéngjìu* |
    | one | CLF:PIECE | considerable | MOD | success, achievement |
    'considerable success'

Similarly, shape classifiers can also be used with varied nouns to denote the physical or conceptual shape of concrete entities and abstract concepts. Example (67) gives the collocations of *tiáo* with various types of nouns.

(67) Collocations of *tiáo* with various types of nouns

a. | *yī* | *tiáo* | *góu* |
    | one | CLF:SLENDER | dog |
    'one dog'

b. | *yī* | *tiáo* | *gēbo* |
    | one | CLF:SLENDER | arm |
    'one arm'

c. | *yī* | *tiáo* | *xiàn* |
    | one | CLF:SLENDER | line |
    'one line'

d. | *yī* | *tiáo* | *zhōngjiān* | *lùxiàn* |
    | one | CLF:SLENDER | middle | road.line, route |
    'the middle way, being neutral'

As shown in example (67), it is clear that *tiáo* 'CLF:SLENDER' is used with the first three nouns in (a-c) based on the long and slender shape of their referents. In (d), it is extended to the referent of *lùxiàn* 'road.line, route' to denote the conceptually long and slender shape of a way or a route. The use of these specific entity classifiers with abstract nouns can also show that they can be metaphorically mapped to abstract nouns to express conceptual properties of noun referents, as discussed in §3.4.2.

In summary, Chinese numeral classifiers can be ranked in terms of the individuation hierarchy. General classifiers, including *zhī* 'CLF:SINGLE', are more likely to be used as default classifiers to show different degrees of individuation, by assigning *gè* and *zhī* to more individuated entities while assigning *zhǒng* to less individuated entities. Specific entity classifiers and event classifiers, on the other hand, tend to be used with more individuated entities by denoting their physical or conceptual features. Therefore, they tend to complement general classifiers *gè* and *zhī* and rank at the top or in the middle of the hierarchy.

5.3.5. Variation in the use of Chinese numeral classifiers

While the distribution of numeral classifiers on the individuation hierarchy can show their semantic correlation with noun referents, there is variation in the use of numeral classifiers and thus nouns can be upgraded or downgraded on the hierarchy. Such variation occurs typically with the general classifiers. As mentioned above, the general kind classifier *zhǒng* is typical for uncountable nouns. When used with countable abstracts of unspecified features, nouns collocated with *zhǒng* can usually be used with the general classifier *gè*. However, concrete countable nouns used with *zhǒng* and *gè* may refer to different entities ranked on different levels on the individuation hierarchy. They are ordered higher on the hierarchy when used with *gè* by being more distinguishable as specific individual referents. However, when concrete countable nouns are collocated with *zhǒng*, they refer to a type of referents or all instances or members in one type, and thus they are less specified in terms of individuation and ranked lower on the hierarchy.

(68) The use of *zhǒng* to refer to one kind or type of referents

   a. *yī*   *zhǒng*   *zhíwù*
       one   CLF:KIND, GENERAL   plant
       'one plant'

   b. *yī*   *zhǒng*   *yìng*   *miànbǐng*
       one   CLF:KIND, GENERAL   hard   pastry
       'one type of hardtack'

   c. *yī*   *zhǒng*   *lǜwǎng*
       one   CLF:KIND, GENERAL   mesh.filter
       'mesh filters'

As shown in example (68), *zhíwù* 'plant' in (a), *yìng miànbǐng* 'hard.pastry, hardtack' in (b), and *lǜwǎng* 'mesh filter' in (c) refer to three types of referents, as they cooccur with *zhǒng*. *Zhíwù* 'plant' in (a) is a hypernym of such nouns as 'tree', 'grass', and 'flower', and the phrase is directly translated into 'one plant'. However, *yìng miànbǐng* 'hard.pastry, hardtack' in (b) is translated into 'one type of hardtack', as it is used with *zhǒng* and refers to an unspecified type of hardtack, e.g., in different shapes or of different flavours. Similarly, example (c) is translated as 'mesh filters' in plural forms in English, for *lǜwǎng* 'mesh filter' is used with the general kind classifier and thus refers to all instances of such kind of mesh filters. Other specific kind classifiers are also used to refer to a type of referents or all instances in one type.

Another example of variation in the use of general classifiers concerns using the general classifier *gè* with nouns for animals. While *gè* is usually collocated with human nouns among animate nouns, as discussed in §5.3.3, it is also used with a small proportion (6.81%) of animal nouns. These nouns can be regarded as being upgraded on the individuation hierarchy. For example, *gǒu* 'dog', *hóu* 'monkey', *jīnglíng* 'genie, elf', *lǎotāo* 'epicure, glutton', and *shòurén* 'beast, orc', are personified as their referents are usually regarded as companions for human beings or as analogous to human beings in appearance or in some features. Other nouns, e.g., *dàxiàng* 'elephant', *dàixióng* 'wombat', and *guàiwù* 'monster', may be more distinguishable in their size, movement or other properties, and they can be upgraded on the hierarchy as well. The results echo Frankowsky and Ke (2016: 64-65) in that *gè* can be used with animal nouns whose referents are either very close to humans or distinguishable by being 'very inhuman'.

Among the collocations with human nouns, *gè* and *wèi* 'CLF:INDIVIDUAL, RESPECT' can be used to show a different social status of human referents, as discussed in §5.3.3. Variation in the use of numeral classifiers can be reflected in the switch of the two

numeral classifiers, e.g., using *gè* instead of *wèi* with *lǎorén* 'elder' but *wèi* instead of *gè* with *háizǐ* 'child'. Since such examples need to be examined in terms of their discourse uses, the variation concerning *gè* and *wèi* and related functions will be analysed in §7.5.

Some specific classifiers may contribute new meanings to nouns when their referents have a different rank on the individuation hierarchy. For example, a small proportion of uncountable nouns are used with specific entity classifiers (23.90%), typically *céng* 'CLF:LAYER' (15.72%). While the shape classifier *céng* 'CLF:LAYER' denotes one dimension when used with countable nouns, it attributes the property of extendedness to referents of such uncountable nouns as *zhīfáng* 'fat' and *guī* 'silicon' in example (69).

(69) The use of entity classifiers with uncountable nouns
    a. *yī*   *céng*   *zhīfáng*
        one   CLF:LAYER   fat
        'a layer of fat'
    b. *yī*   *céng*   *gāochún*   *guī*
        one   CLF:LAYER   pure   silicon
        'a thin pure silicon layer'

Similarly, event classifiers can coerce an event reading when collocated with non-event nouns. For example, *fàn* 'rice' generally refers to cooked grain and *diànhuà* 'telephone' refers to a telephone. However, they can refer to a meal and a phone call, respectively, when they cooccur with the event classifiers *dùn* 'CLF:SPELL, SESSION' and *cì* 'CLF:TIME', as shown in example (70).

(70) The use of event classifiers with concrete nouns
    a. *yī*   *dùn*   *fàn*
        one   CLF:SPELL, SESSION   rice
        'a meal'
    b. *yī*   *cì*   *běndì*   *diànhuà*
        one   CLF:TIME   local   telephone
        'a local phone call'

In summary, variation in the use of Chinese numeral classifiers may involve the reassignment of nouns on the order of the individuation hierarchy. Numeral classifiers can also contribute new meanings to nouns by applying specific entity classifiers to uncountable nouns and event classifiers to non-event nouns. The variation can also be interpreted in

terms of the semantic functions concerning "ascribing properties to referents" and the discourse function of "representation of referents", as discussed in §3.3.

5.4. Collocations of numeral classifiers in Corpus 2

This section is devoted to the collocations of Chinese numeral classifiers in the context of adjectives in Corpus 2. Their collocations with adjectives will be examined first in §5.4.1, and the collocations with nouns will be discussed in §5.4.2.

5.4.1. Collocations with adjectives

Adjectives pre-modifying numeral classifiers usually denote size or shape. As shown in Table 24, only nine adjectives are found to be modifiers of numeral classifiers in 523 classifier phrases. *Xiǎo* 'small, little' is the most frequent one, constituting 76.86% of all the cooccurrences, while its antonym *dà* 'big, large' (12.62%) is the second on the list. This pair of adjectives makes up 88.48% of all the collocations with numeral classifiers. In contrast, the other seven adjectives, sharing the remaining 12.52%, are related to shape. Except for *zhěng* 'whole, full, entire', which denotes the wholeness of the referent, the others are related to length and thickness, as in *cháng* 'long' vs. *duǎn* 'short', and *hòu* 'thick' vs. *báo* 'thin', *xì* 'slender, thin', and *zhǎi* 'narrow'.

Table 24. The adjectives collocated with Chinese numeral classifiers

specific entity classifier	gloss	occurrence	
		#	%
xiǎo	'small, little'	402	76.86
dà	'big, large'	66	12.62
cháng	'long'	25	4.78
báo	'thin'	20	3.82
zhǎi	'narrow'	4	0.76
hòu	'thick'	3	0.57
xì	'slender, thin'	1	0.19
zhěng	'whole, full, entire'	1	0.19
duǎn	'short'	1	0.19
total		523	

The adjectives *xiǎo* 'small, little' and *dà* 'big, large' can be applied to almost all numeral classifiers in Corpus 2. As indicated in example (71), the two adjectives are used with the shape classifier *kuài* 'CLF:LUMP.LIKE' in (a), the entity classifier denoting other salient features *jiān* 'CLF:ROOM' in (b), the general classifier *gè* in (c), the event classifier *cháng* 'CLF:VENUE' in (d), and the kind classifier *lèi* 'CLF:CATEGORY' in (e). While the referents of the first four numeral classifiers can be described as big or small based on their physical or temporal bounds, *lèi* 'CLF:CATEGORY' is the only kind classifier whose referents that can be measured in terms of size or scale. The results show that numeral classifiers pre-modified by adjectives tend to denote discrete units of entities and events or sometimes categories.

(71) The collocation of numeral classifiers with adjectives *xiǎo* and *dà*
    a. *yī*   *xiǎo*   *kuài*   *miànbāo*
      one   small, little CLF:LUMP.LIKE   bread
      'a very little strip of bread'
    b. *yī*   *xiǎo*   *jiān*   *wòshì*
      one   small, little CLF:ROOM   bedroom
      'a small space'
    c. *yī*   *dà*   *gè*   *shēngrì*   *dàngāo*
      one   big, large CLF:GENERAL   birthday   cake
      'a big birthday cake'
    d. *yī*   *dà*   *cháng*   *jià*
      one   big, large CLF:VENUE   fight
      'a big row'
    e. *yī*   *dà*   *lèi*   *chángjiàn* *jíbìng*
      one   big, large CLF:CATEGORY common   disease
      'a clinical symptom complex'

Other adjectives are more restricted in the collocations with numeral classifiers. For example, the adjective *zhěng* 'whole, full, entire' can only be used with numeral classifiers that denote discrete units, and thus, it can replace *dà* and *xiǎo* in all examples in (71), except for *dà* with *lèi* 'CLF:CATEGORY' in (e). Other adjectives are more restricted by numeral classifiers in terms of shape. For example, *cháng* 'long' is more likely to be used with the numeral classifier *tiáo* 'CLF:SLENDER', while *hòu* 'thick' and *báo* 'thin' tend to cooccur with the numeral classifier *céng* 'CLF:LAYER'.

    The presence of adjectives, on the other hand, restricts the choice of numeral classifiers. As discussed in §5.2, the numeral classifiers pre-modified by adjectives are predominantly specific entity classifiers (96.7%), typically shape classifiers (92.7%), e.g.,

kuài 'CLF:LUMP-LIKE' (52.9%) and *piàn* 'CLF:FLAT/THIN' (17.6%). The general classifier *gè* (0.8%) rarely appears with adjectives, and the general kind classifier *zhǒng* cannot be used with adjectives. What should also be noted is that specific classifiers for humans cannot be preceded with adjectives, and specific classifiers for animals, e.g. *zhī* 'CLF:SIN-GLE' and *toú* 'CLF:HEAD' do not occur in the corpus.

The two-way restrictions of adjectives and numeral classifiers can be accounted for by their semantic preference for the shared features (Sinclair 2004: 142). The co-selection of adjectives and numeral classifiers requires that numeral classifiers be delimitable in terms of the semantic features expressed by adjectives, or vice versa. Therefore, the general kind classifier *zhǒng* cannot be modified by adjectives, as it is unspecified and non-discrete. Specific classifiers for humans, e.g., *wèi* 'CLF:INDIVIDUAL, RESPECT' and *míng* 'CLF:IDENTITY', cannot be modified by adjectives either, since their semantic features cannot be shared by the above adjectives. The general classifier *gè* and other specific classifiers based on other features than shape, e.g., for animals and plants, are less likely to be preceded by adjectives since *gè* is unspecified and the others are less closely associated with shape and size. In contrast, shape classifiers are more likely to cooccur with adjectives as both of them have a quantity reading based on their shared features of shape and size.

To conclude, there are two-way restrictions in the choice of adjectives and numeral classifiers based on their semantic preference. Numeral classifiers in the context of adjectives are more likely to denote shape, while the adjectives used with numeral classifiers tend to denote size and shape. The quantity reading of shape and size is more typical of measure words, which implies that numeral classifiers in the context of adjectives may cooccur with a wider range of nouns and be more likely to be translated into measure words in English. Their collocations with nouns will be examined in the next section, and the comparison of the direct translations of numeral classifiers in the context with and without adjectives which will be examined in Chapter 7.

5.4.2. Collocations with nouns

The distribution of numeral classifiers in Corpus 2 shows that they are rarely collocated with nouns for humans and animals in the context of adjectives. As shown in Table 25,

only one animate as well as human noun is found in the corpus, and it is a collective noun for humans, *rénqún* 'crowd'.

Table 25. The distribution of numeral classifiers in Corpus 2

semantic group	numeral classifiers	
	#	%
+human	1	0.19
-human	522	99.81
±human	0	0
+animate	1	0.19
-animate	522	99.81
±animate	0	0
+concrete	478	91.40
-concrete	27	5.16
±concrete	18	3.44
+count	384	73.42
-count	60	11.47
±count	80	15.30
total	523	

As regards the semantic opposition concerning concreteness, numeral classifiers in the context of adjectives are predominantly used with concrete nouns (91.40%), while far lower proportion of abstract nouns (5.16%) occur with numeral classifiers in the context of adjectives in Corpus 2, compared with the proportion of abstract nouns occurring in Corpus 1 (36.73%, see Table 13). Similarly, numeral classifiers occur far more frequently with countable nouns in Corpus 2 (73.42%). However, compared with the percentage (2.37%, see Table 13) of numeral classifiers cooccurring with uncountable nouns in Corpus 1, a higher percentage (11.47%) of numeral classifiers are used with uncountable nouns in Corpus 2. As numeral classifiers in the context of adjectives are predominantly entity classifiers (see §5.2), this section will focus on the collocations of entity classifiers in association with two semantic oppositions concerning concreteness and countability.

Entity classifiers are predominantly used with concrete nouns. As shown in Table 26, 92.94% of them are used with concrete nouns, while the remaining part (3.92% and 3.14% respectively) are collocated with abstract nouns and nouns unspecified in terms of concreteness.

Table 26. The distribution of entity classifiers in terms of concreteness in Corpus 2

semantic oppositions	#	%
+concrete	474	92.94
-concrete	20	3.92
±concrete	16	3.14
total2	510	

Among concrete nouns, entity classifiers cooccur predominantly with inanimates, as mentioned above. The entity classifiers are less varied in the context of adjectives than those not pre-modified by adjectives in Corpus 1. However, shape classifiers in the context are collocated with more varied inanimate nouns. As shown in example (72), *tiáo* 'CLF:SLENDER' can be used with *gōnglù* 'road' in (a) and (b) no matter whether the adjective *cháng* 'long' is present. However, it cannot be collocated with nouns in (c)-(e) if the adjectives are absent. Without the presence of adjectives, *tiānkōng* 'sky' is more likely to cooccur with *piàn* 'CLF:FLAT/THIN' based on the feature of extendedness. Similarly, *lùdì* 'land' should be used with *piàn* 'CLF:FLAT/THIN' or *kuài* 'CLF:LUMP.LIKE'. As to *tiān'é'róng* 'velvet', it is usually collocated with the general kind classifier, as it is uncountable.

(72) The use of *tiáo* with concrete inanimates

a. *yī*   *tiáo*   *gāosù*   *gōnglù*
    one   CLF:SLENDER   high.speed   road
    'a highway'

b. *yī*   *cháng*   *tiáo*   *gāosù*   *gōnglù*
    one   long   CLF:SLENDER   high.speed   road
    'a stretch of highway'

c. *yī*   *cháng*   *tiáo*   *fánxīngmìbù*   *de*   *tiānkōng*
    one   long   CLF:SLENDER   starry   MOD   sky
    'a strip of starry sky'

d. *yī*   *xiǎo*   *tiáo*   *xiácháng*   *lùdì*
    one   small, little   CLF:SLENDER   long.and.narrow   land
    'a little strip of land'

e. *yī*   *cháng*   *tiáo*   *tiān'é'róng*
    one   long   CLF:SLENDER   velvet
    'a strip of velvet'

Among the limited number of abstract nouns, most are used with entity classifiers (20 out 27). While non-observable abstract nouns are predominantly used with the general classifiers *gè* (63.7%) and *zhǒng* (23.8%) in Corpus 1, as discussed in §5.3.3, in the context of adjectives in Corpus 2, none of them appears with the two general classifiers. Instead,

they are more likely to cooccur with shape classifiers, typically *kuài* 'CLF:LUMP.LIKE' (25.9%) and *duàn* 'CLF: SEGMENT' (25.9%). As shown in example (73), *kuài* 'CLF:LUMP.LIKE' is collocated with such nouns as *gōngzuò* 'job, work', *gǔquán* 'share.holding' and *dàimǎ* 'code', which are more likely to cooccur with the general classifier *gè* in Corpus 1.

(73) The use of *kuài* with abstract nouns
   a. *yī*   *xiǎo*   *kuài*   *gōngzuò*
      one   small, little   CLF:LUMP.LIKE   job, work
      'this narrow realm of practice'
   b. *yī*   *xiǎo*   *kuài*   *gǔquán*
      one   small, little   CLF:LUMP.LIKE   share.holding
      'blocks of (PPCW) shares'
   c. *yī*   *xiǎo*   *kuài*   *dàimǎ*
      one   small, little   CLF:LUMP.LIKE   code
      'a little procedure'

Entity classifiers are also mainly used with countable nouns. As shown in Table 27, 73.53% of them occur with countable nouns, in opposition to 11.37% with uncountable nouns and 15.10% with nouns annotated as [±count]. Among the countable nouns, most of them are concrete.

Table 27. The distribution of entity classifiers in Corpus 2

semantic oppositions	#	%
+count	375	73.53
-count	58	11.37
±count	77	15.10
total	510	

It should be noted that uncountable nouns are collocated far more frequently with entity classifiers (58 out 60) in the context of adjectives in Corpus 2, compared with the percentage of uncountable nouns used with entity classifiers (42 out of 159, see Table 13) in Corpus 1. As the unspecified general kind classifier *zhǒng* cannot be modified by adjectives of shape and size, more specific shape classifiers are used with uncountable nouns. As illustrated in example (74), such uncountable nouns as *shíwù* 'food' in (a), *sùliào* 'plastic' in (b), *cáiliào* 'material' in (c), and *zhīwù* 'fabric, textile' in (d) appear with shape classifiers, although their default numeral classifier is *zhǒng* when they are not preceded with adjectives.

(74) The use of specific classifiers with uncountable nouns in the context of adjectives

    a. *yī*   *xiǎo*   *kuài*   *shíwù*
       one   small, little   CLF:LUMP.LIKE   food
       'a morsel of food'

    b. *yī*   *báo*   *céng*   *sùliào*
       one   thin   CLF:LAYER   plastic
       'a layer of plastic'

    c. *yī*   *xiǎo*   *piàn*   *cáiliào*
       one   small, little   CLF:FLAT/THIN   material
       'a small piece of material'

    d. *yī*   *cháng*   *tiáo*   *miánbù*   *huò*   *yàmábù de*   *zhīwù*
       one   long   CLF:SLENDER   cotton   or   linen   MOD fabric, textile
       'a long piece of cotton or linen cloth'

The results also show that shape classifiers pre-modified with adjectives cooccur with a wider scope of nouns without necessarily denoting their inherent properties. As indicated example (74), such nouns as *shíwù* 'food', *sùliào* 'plastic', *cáiliào* 'material', and *zhīwù* 'fabric, textile' are vague in terms of their physical bounds and unspecified with regard to shape and size. However, with the presence of adjectives, they can refer to more specific referents, as shape classifiers pre-modified with adjectives denote more specific size and shape of their referents and thus contribute new meanings to noun phrases.

    In conclusion, numeral classifiers pre-modified with an adjective are more likely to be used with inanimates and tend to be more specified shape classifiers. Furthermore, shape classifiers in the context of adjectives are more like measure words by being used with a wider range of nouns without necessarily denoting the features of noun referents. In this context, they contribute new meanings concerning size and shape to noun phrases and refer to more specific referents.

5.5. Discussion and concluding remarks

This section will first give a brief summary of the findings in this chapter, and then more attention will be devoted to the semantic contribution of Chinese numeral classifiers and other elements to noun phrases. The above sections have shown the frequency and collocations of Chinese numeral classifiers. Generally speaking, general classifiers are used far more frequently than specific classifiers, which tend to express more explicit

properties. In the context of an adjective, however, specific classifiers, typically shape classifiers, are used far more frequently than general classifiers.

Chinese numeral classifiers can be classified based on the individuation hierarchy (Fletcher 1987; Audring 2006: 94). While the three general classifiers, *gè* 'CLF:GENERAL', *zhī* 'CLF:SINGLE', and *zhǒng* 'CLF:KIND, GENERAL' are used as default classifiers to show different degrees of individuation, specific classifiers are used to show more salient semantic properties of different types of nouns. There is also variation in the use of Chinese numeral classifiers. The general classifier *gè* appear not only as a default classifier for human nouns but also with nouns for animals. Similarly, the general kind classifier *zhǒng* is used not only as a default classifier for uncountable nouns but also with countable nouns. Among specific classifiers, some entity classifiers can be used with uncountable nouns, and a few event classifiers also cooccur with non-event nouns. Shape classifiers pre-modified with an adjective are more likely to cooccur with a wider range of nouns without necessarily being semantically consistent with the nouns, as shown in (c)-(e) in example (72).

Numeral classifiers and other constituents can thus contribute to the semantics of the noun phrases in several ways. General classifiers can be used with broad categories of nouns to express such properties as humanness, animacy, and boundedness. On the other hand, nouns cooccurring with general classifiers tend to be more specific and refer to more specific referents, which can account for the high frequency of general classifiers and their application to more varied nouns. As to specific classifiers, they tend to be semantically correlated with their head nouns and specify properties of noun referents. Some specific classifiers, e.g., *tiáo* 'CLF:SLENDER', specify a physical or conceptual property, as with the concrete noun *lù* 'road' or the abstract noun *lùxiàn* 'road.line, route'. Other specific classifiers, e.g., *běn* 'CLF:BOOK', overlap with nouns in terms of reference, as with *shū* 'book'.

Regarding variation in the use of numeral classifiers, numeral classifiers can contribute new meanings to noun phrases. For example, when *diànhuà* 'telephone' is used with the event classifier *cì* 'CLF:TIME', the referent should be interpreted as a phone call rather than a telephone. Shape classifiers pre-modified by an adjective can contribute new meanings to noun phrases. While shape classifiers denote features related to shape, dimension or extendedness of referents and help speakers to distinguish the referents, they also have a quantity reading when used with adjectives of size and shape. Therefore, the

semantics of the noun phrase in this context should be interpreted based on the meaning of all the elements, as shown in example (75).

(75) The semantic contribution of numeral classifiers, adjectives and nouns

    a) *yī*   *cháng*   *tiáo*   *tiānkōng*
       one   long   CLF:SLENDER   sky
       'a long stretch of the sky'

    b) *yī*   *xiǎo*   *kuài*   *píngguǒ*
       one   small, little CLF:LUMP.LIKE   apple
       'a bit of apple'

While *tiānkōng* 'sky' in (a) refers to the sky, the adjective *cháng* 'long' and the shape classifier *tiáo* 'CLF:SLENDER' both indicate that the long and narrow shape of the referent is a part of the sky instead of the sky in general. In (b), *kuài* 'CLF:LUMP.LIKE' refers to a piece in an apple instead of a complete apple, and the adjective *xiǎo* 'small, little' contributes the meaning related to size to the referent. Thus, the referent of the noun phrase is not an apple, but a small lump-like piece of an apple, as also shown in the English translation. Therefore, numeral classifiers can contribute extra meanings related to both quality and quantity to their referents.

Chapter 6: Chinese numeral classifiers in translation: semantic functions

6.1. Introduction

Numeral classifiers are one of the features that distinguish Chinese from English, or more generally, a classifier language and a non-classifier language. While there are no numeral classifiers in English, it has other forms to express the lexical and grammatical meanings of numeral classifiers and to represent their related functions in translation. This chapter will compare Chinese numeral classifiers with their equivalent forms in English translation based on their lexical and grammatical meanings and semantic functions. To be more specific, this chapter aims to address the following research questions based on Corpus 1 and Corpus 2:

a) To what extent are Chinese numeral classifiers directly equivalent to English measure words?

b) How are their grammatical meanings involving definiteness and lexical meanings concerning specific properties of referents represented in English, typically when Chinese numeral classifiers are not equivalent to measure words in translation between the two languages?

c) What equivalent forms in English are used to individuate nouns, differentiate referents and ascribe properties to referents?

The translation of numeral classifiers and the numeral *yī* 'one' is closely related to the grammatical meanings of Chinese numeral classifiers and their related semantic function concerning individuation. Therefore, I will address these issues in §6.2. In §6.3, I will focus on the representation of the semantic function involving differentiating

referents in English translation, based on the lexical meanings concerning specific properties of Chinese numeral classifiers. The semantic function of ascribing properties to referents will be dealt with in §6.4, typically based on the variation in the use of Chinese numeral classifiers. Discussion and conclusions will be given in §6.5.

6.2. Individuation of nouns

The obligatory use of Chinese numeral classifiers is attributed to the need to create a semantic unit for quantifying nouns (see §3.2.1 and §3.3.1). Numeral classifiers are usually regarded as complementary to plural marking in non-classifier languages (see §2.4.2). They can be compared with English measure words in translations between the two languages and the individuation can be reflected in the choice of singular and plural forms typically when they are omitted in English translation. Therefore, the direct equivalents of Chinese numeral classifiers in English will be dealt with in §6.2.1, and the translation of the numeral *yī* 'one' will be examined in §6.2.2.

6.2.1. The equivalence of numeral classifiers as measure words

While numeral classifiers tend to be omitted when translated into non-classifier languages, Chinese classifiers can be equivalent to English measure words in pseudo-partitive structures in the translation between the two languages, typically in the context of adjectives. As shown in example (76), *céng* 'CLF:LAYER' in (a) and *kuài* 'CLF:LUMP.LIKE' in (b) are equivalent to the measure words *layer* for the mass noun *sunblock* and *piece* used with the adjective *small* and the concrete noun *cake*, respectively.[25]

(76) The equivalence of Chinese numeral classifiers in pseudo-partitive structures in English
  a. *yī*   *céng*   *fángshàishuāng*
    one   CLF:LAYER   sunblock
    'a layer of sunblock'

---

[25] Examples and their English translations in this chapter are collected from Corpus 1 and Corpus 2 unless otherwise indicated.

b. *yī*   *xiǎo*   *kuài*   *dàngāo*
   one   small   CLF:LUMP.LIKE   cake
   'a small piece of cake'

Chinese numeral classifiers are more likely to be equivalent to English measure words in the context of adjectives. Table 28 shows the proportion of the occurrence of English measures words as equivalence of Chinese numeral classifiers in Corpus 1 and Corpus 2. Total 1 and total 2 refer to the total numbers of numeral classifiers used with their related types of nouns in the two corpora, while total 3 shows the total occurrences of equivalent measure words in English. As shown in the table, 3.06% of numeral classifiers in Corpus 1 are equivalent to English measure words. In contrast, as many as 58.13% of numeral classifiers in the context of adjectives in Corpus 2 are equivalent to English measure words.

Table 28. The equivalence of Chinese numeral classifiers as measure words in English

	Corpus 1			Corpus 2		
	#	%	total 1	#	%	total 2
+concrete	121	3.06	3959	284	59.41	478
-concrete	75	3.05	2461	13	48.15	27
±concrete	9	3.21	280	7	38.89	18
+count	106	1.94	5448	203	52.86	384
-count	33	20.75	159	48	80.00	60
±count	66	6.04	1093	53	67.09	79
total 3	205	3.06	6700	304	58.13	523

The degree to which numeral classifiers are equivalent to measure words in the translation between the two languages depends partly on whether they occur with concrete or abstract nouns. As shown in Table 28, numeral classifiers used with concrete nouns are more frequently equivalent to English measure words (59.41%) in Corpus 2 compared with those collocated with the other two groups of nouns, i.e., abstract and polysemous nouns (38.89%-48.15%). However, little difference is found in the direct translation of numeral classifiers when collocated with concrete or abstract nouns in Corpus 1 (3.05%-3.21%).

Table 28 also shows that the likelihood of the occurrence of English measure words as equivalence of Chinese numeral classifiers is associated with countability. In Corpus 1, a significantly higher proportion of 20.75% of numeral classifiers are equivalent to measure words in English when they cooccur with uncountable nouns, in opposition to 1.94% and 6.04% occurring with countable nouns and nouns annotated with

[±count]. A higher percentage of 80% is also found in the direct translation of numeral classifiers collocated with uncountable nouns in Corpus 2.

In summary, Chinese numeral classifiers can be directly equivalent to Egnlish measure words in the translation between the two languages, typically when numeral classifiers are pre-modified by adjectives or collocate with uncountable nouns. The translation of the specific features denoted by numeral classifiers will be further analysed in §6.3.

6.2.2. The corresponding form of yī 'one' in English

The grammatical features of numeral classifiers concerning definiteness can be shown in the corresponding forms of the numeral *yī* 'one', typically when numeral classifiers are not equivalent to English measure words. This section will compare the numeral *yī* 'one' with determiners in the translation between the two languages in Corpus 1 and Corpus 2 to show the influence of numeral classifiers on the interpretation of the numeral *yī* 'one' and its representation in English translation. Table 29 shows the corresponding forms of the numeral *yī* 'one' in English.

Table 29. The corresponding forms of the numeral *yī* 'one' in English translation

English equivalent of *yī* 'one'	Corpus 1		Corpus 2	
	#	%	#	%
indefinite article	5546	82.78	422	80.69
other indefinite determiner	477	6.27	58	11.09
definite article	351	5.24	23	4.39
other definite determiner	67	1.00	11	2.10
numeral *one*	319	4.76	9	1.72
total	6700		523	

As shown in Table 29, *yī* 'one' is rarely corresponding to *one* (4.76% in Corpus 1 and 1.72% in Corpus 2), although the direct equivalent of the numeral *yī* is *one* in English. Instead, the numeral tends to be comparable with articles and other determiners. It is predominantly corresponding to the indefinite articles *a* and *an* (82.78% in Corpus 1 and 80.69% in Corpus 2), as shown in (a) in example (77). Furthermore, its equivalent in English can also be indefinite determiners (6.27% in Corpus 1 and 11.09% in Corpus 2), e.g., *any* in (b) and *some* in (c) in example (77).

(77) Indefinite readings of *yī* 'one' shown in the indefinite articles and determiners in English
    a. *yī*   *gè*   *guānjiàn*   *yīnsù*
       one   CLF:GENERAL   key   factor
       'a key factor'
    b. *yī*   *bǐng*   *bǎojiàn*
       one   CLF:STIPE   sword
       'any sword'
    c. *yī*   *zhī*   *xúnliáng*   *de*   *dà*   *tùzi*
       one   CLF:SINGLE   docile   MOD   big   rabbit
       'some harmless rodent'

The indefinite reading of *yī* used with numeral classifiers is not only shown in its equivalent indefinite articles and other indefinite determiners in English. It is also shown in the form of bare nouns with both the numeral and numeral classifier omitted or in the plural form of a noun used in English. Example (78) shows the indefinite readings of *yī* shown in bare nouns or nouns in plural forms in English.

(78) Indefinite readings of *yī* 'one' shown in bare nouns and nouns in plural forms in English
    a. *yī*   *zhǒng*   *àiqíng*
       one   CLF:KIND, GENERAL   love
       'love'
    b. *yī*   *piàn*   *liánmiánbùjué*   *de*   *shān*
       one   CLF:FLAT/THIN   endless   MOD   mountain
       'endless mountains'
    c. *yī*   *gè*   *guāncházhe*
       yī   CLF:GENERAL   observer
       'observers'
    d. *yī*   *zhǒng*   *tuōcí*
       one   CLF:KIND, GENERAL   excuse, reason
       'all kinds of reasons'

As shown in example (78), *yī* is omitted together with the numeral classifiers in the English translations. In (a), *àiqíng* 'love', an uncountable noun collocated with the classifier *zhǒng*, is equivalent to bare noun *love* in English, an uncountable noun. However, the nouns in (b) and (c) are both represented in their equivalents in plural forms. The numeral classifier *piàn* 'CLF:FLAT/THIN' coerces a plural reading of the referent in (c) by attributing the property of extendedness of a range of mountains, as discussed in §5.3.5. In (c), the equivalence of *guāncházhe* 'observer' with 'observers' can be attributed to the indefinite reading of *yī* used with a numeral classifier. *Yī* 'one' used with numeral classifiers in

Chinese can be analogous with the indefinite determiners *every* or *each* in English, which occurs with a singular noun but have a plural connotation. Since such indefinite and plural readings of *yī* 'one' used with numeral classifiers should be analysed based on their context, further examples will be discussed in §7.4.1. Finally, the indefinite reading of *yī* with numeral classifiers can also be shown when they are corresponding to such determiners as *all* and *a few* with plural nouns in English, as shown in (d).

    A small proportion of noun phrases in English translation include definite determiners, including the definite article *the* (5.24% in Corpus 1 and 4.39% in Corpus 2) and such definite determiners as demonstratives, e.g., *this* and *that*, and possessive pronouns, e.g., *his* and *their,* and a possessive form of indefinite pronouns, e.g., *somebody's* (1.00% in Corpus 1 and 2.10% in Corpus 2). Example (79) shows the definite readings of *yī* used with Chinese numeral classifiers.

(79) Definite readings of *yī* used with Chinese numeral classifiers
    a. *yī*   *gè*   *shìjiè*
       one   CLF:GENERAL   world
       'the world'
    b. *yī*   *báo*   *céng*   *shuǐ*
       one   thin   CLF:LAYER   water
       'this film of water'
    c. *yī*   *zhǒng*   *shúxī*   *de*   *bùshì*
       one   CLF:KIND, GENERAL familiar   MOD   discomfort
       'that familiar (claggy-mouthed) discomfort'
    d. *yī*   *gè*   *shìpín*
       one   CLF:GENERAL   video
       'their video'
    e. *yī*   *gè*   *jiānbǎng*
       yī   CLF:GENERAL   shoulder
       'somebody's shoulder'

As shown in example (79), *yī* is corresponding to the definite article *the* in (a), demonstratives *this* in (b) and *that* in (c), the possessive pronoun *their* in (d) and the possessive form of the indefinite pronoun *somebody's* in (e). The examples show that the presence of numeral classifiers can influence the interpretation of *yī* 'one' to express definiteness and indefiniteness, other than the specific quantity. The definiteness of numeral classifiers will also be further analysed in §7.4.1, based on discourse data.

    The above two sections have shown that some Chinese numeral classifiers can be equivalent to English measure words in the translation between the two languages,

typically in the context of uncountable nouns and adjectives. When they are used with uncountable nouns, they tend to create semantic units to facilitate quantification. In the context of countable nouns, the numeral classifiers equivalent to English measure words are likely to contribute additional information concerning quantity or quality to noun phrases. When there are no measure words in English, the function of numeral classifiers concerning the individuation of nouns can be reflected in the translation of the definite and indefinite reading of *yī* 'one' coerced by the presence of numeral classifiers, as *yī* 'one' can be shown in articles and determiners, as well as bare nouns or nouns in plural forms.

6.3. Differentiating referents

Chinese numeral classifiers can denote specific properties related to humanness, animacy and shape, as shown in §3.2.2. They can also be ranked in the individuation hierarchy (see Table 23) based on the degree of countability and their semantic correlation with nouns as shown in §5.3.4. This section will examine how specific properties expressed by numeral classifiers are shown in English translation and how the related function concerning differentiating referents is reflected in English. The English translation of the specific properties of numeral classifiers will first be investigated based on the English equivalents of numeral classifiers in the two corpora in the context with and without adjectives in §6.3.1. In §6.3.2, I will examine the English equivalents of nouns and the representation of the function of differentiating referents.

6.3.1. Specific properties reflected in the English equivalence of numeral classifiers

It has been shown in §6.2.1 that numeral classifiers can be directly equivalent to English measure words to quantify nouns. This section will examine English measure words as equivalence of numeral classifiers to show to what degree and how they express the specific properties denoted by numeral classifiers. Chinese numeral classifiers will be compared with English measure words, in terms of their specific properties in the context

without adjectives in Corpus 1 in §6.3.1.1, and in the context with adjectives in Corpus 2 in §6.3.1.2.

6.3.1.1. In the context without adjectives in Corpus 1

In the context without adjectives, specific entity classifiers that denote such features as humanness, animacy and physical properties and the general kind classifier *zhǒng* are more likely to be equivalent to English measure words. As shown in Table 30, among the 205 English equivalents, 49.76% are equivalent to entity classifiers (including 9.76% for the general classifier *gè* and 40% for specific entity classifiers), and another 47.81% (including 40.79% for the general kind classifier *zhǒng* and 7.3% for specific entity classifiers) are equivalent to kind classifiers, while the remaining 2.44% are equivalent to event classifiers.

Table 30. The distribution of English measure words as equivalence to different types of Chinese numeral classifiers in Corpus 1

type of numeral classifiers		direct translation of numeral classifiers	
		#	%
entity	general	20	9.76
	specific	82	40.00
event		5	2.44
kind	general	83	40.49
	specific	15	7.32
total		205	

Entity classifiers are predominantly used with countable nouns, and therefore, more entity classifiers used with countable nouns are equivalent to English measure words. As shown in Table 31, among the 102 measure words, 67 of them (65.69%) are equivalent to entity classifiers collocated with countable nouns. However, only 14 entity classifiers (13.73%) cooccurring with uncountable nouns are equivalent to measure words.

Table 31. The distribution of English measure words as equivalence to entity classifiers in Corpus 1

	direct translation of entity classifiers	
	#	%
+count	67	65.69
-count	14	13.73
±count	21	20.59
total	102	

Several reasons can be attributed to the higher proportion of the occurrence of English measure words as equivalence to entity classifiers used with countable nouns. First, compared with 4777 entity classifiers used with countable nouns, only 42 entity classifiers are collocated with uncountable nouns (see Table 13) in Corpus 1. Among the 42 entity classifiers, 14 are equivalent to measure words in English. Furthermore, some numeral classifiers may have a quantity reading when used with countable nouns. In this case, measure words are expected as their English equivalents. As shown in example (80), *dàngāo* 'cake' used with the general classifier *gè* in (a) refers to a cake, while *kuài* 'CLF:LUMP.LIKE' in (b) implies a quantity, and therefore, *kuài* used in such context is comparable with the measure word *piece* in English.

(80) English translation of specific properties indicated by numeral classifiers
    a. *yī*   *gè*   *dàngāo*
        one   CLF:GENERAL   cake
        'a cake'
    b. *yī*   *kuài*   *dàngāo*
        one   CLF:LUMP.LIKE   cake
        'a piece of cake'

Entity classifiers are more likely to be equivalent to such measure words as piece, kind, *bit*, *sort* and *type*. For example, *piece* is equivalent to various entity classifiers, including the general classifier and specific entity classifiers. As shown in example (81), such entity classifiers include the general classifier *gè* in (a), less specified entity classifiers, e.g., *jiàn* 'CLF:PIECE' in (b) and *xiàng* 'CLF:ITEM, PROJECT' in (c), and specific entity classifiers denoting shape, e.g., *kuài* 'CLF:LUMP.LIKE' in (d), *duàn* 'CLF:SEGMENT' in (e), and *tiáo* 'CLF:SLENDER' in (f), and specific entity classifiers denoting other salient features, e.g., *bù* 'CLF:DEMO' in (g). Furthermore, *piece* is also collocated with nouns ranging from concrete nouns as in examples (a)-(d) to abstract nouns as in (e)-(g).

(81) Chinese numeral classifiers equivalent to the measure word *piece* in Corpus 1
    a. *yī*   *gè*   *xiǎo*   *zhǐtiáo*
        one   CLF:GENERAL   *small*   paper.slip
        'a piece of paper'
    b. *yī*   *jiàn*   *suíshēn*   *xíngli*
        one   CLF:PIECE   carry-on   luggage
        'a piece of carry-on luggage'
    c. *yī*   *xiàng*   *jìshù*
        one   CLF:ITEM, PROJECT   technology
        'a piece of technology'

d. *yī*   *kuài*   *dàngāo*
   one   CLF:LUMP.LIKE   cake
   'a piece of cake'

e. *yī*   *duàn*   *qǔzǐ*
   one   CLF:SEGMENT   music
   'a piece of music

f. *yī*   *tiáo*   *hěnhǎo*   *de*   *jiànyì*
   one   CLF:SLENDER   very good   MOD   suggestion, advice
   'a piece of good advice'

g. *yī*   *bù*   *zhěnmì*   *xiángshí*   *de*   *dàzuò*
   one   CLF:DEMO   thorough   detailed   MOD   work
   'a detailed and thorough piece of work'

On the other hand, entity classifiers are also equivalent to measure words that indicate some features of their referents. The most typical example is *céng* 'CLF:LAYER', equivalent to *layer*, *coat*, *cloak*, *mantle* and *veneer* that indicate a thin cover of referents of such nouns as *shadow*, *clay*, *snow*, *frost* and *mud*, as indicated in example (82).

(82) The direct translations of *céng*

a. *yī*   *céng*   *yīnyǐng*
   one   CLF:LAYER   shadow
   'a layer of shadow'

b. *yī*   *céng*   *hòuhòu*   *de*   *lùní*
   one   CLF:LAYER   thick   MOD   clay
   'a thick coat of clay'

c. *yī*   *céng*   *yínzhuāng*
   one   CLF:LAYER   silver.dress, snow
   'a cloak of snow'

d. *yī*   *céng*   *báishuāng*
   one   CLF:LAYER   frost
   'a white mantle of frost'

e. *yī*   *céng*   *báobao*   *de*   *hēisè*   *ruǎnní*
   one   CLF:LAYER   thin   MOD   black   mud
   'a thin veneer of black mud'

Another entity classifier also likely to be equivalent to measure words is *piàn* 'CLF:FLAT/THIN'. It can be translated into *sheet* when used with nouns referring to paper, as indicated in (a) in example (83). However, it can also denote extendedness of referents and is thus comparable with such measure words as *cluster* in (b), *stretch* in (c), and *expanse* in (d) to show different ways of extendedness of the referents of *buildings*, *mountains* and *rays*, respectively, and it is equivalent to *flurry* in (d) to denote the many thanks.

(83) The direct translations of *piàn*

    a. *yī*   *piàn*   *bǎnzhǐ*
       one   CLF:FLAT/THIN   paperboard
       'a sheet of paper'

    b. *yī*   *piàn*   *jiànzhù*
       one   CLF:FLAT/THIN   building, architecture
       'a cluster of buildings'

    c. *yī*   *piàn*   *shānqū*
       one   CLF:FLAT/THIN   mountainous area
       'one stretch of mountains'

    d. *yī*   *piàn*   *jīnguāng*
       one   CLF:FLAT/THIN   golden.ray
       'an expanse of golden rays'

    e. *yī*   *piàn*   *xièxiè*
       one   CLF:FLAT/THIN   thank
       'a flurry of thanks'

What should be noted is that the measure words in English can be equivalent to nouns or modifiers of nouns in Chinese instead of numeral classifiers. As shown in example (84), none of the three measure words in English are equivalents of the general classifier *gè*. Instead, the measure word *stretch* in (a) is translated based on the noun *shānqū* 'mountain.area', which refers to a stretch of mountains, *group* in (b) is translated based on the collective noun *jiězǔ* 'solution.group', and *trace* in (c) is translated based on *guāngmáng* 'light.ray/trace'.

(84) English measure words as equivalence to Chinese nouns or modifiers in Corpus 1

    a. *yī*   *gè*   *shānqū*
       one   CLF:GENERAL   moutain.area
       'a stretch of mountains'

    b. *yī*   *gè*   *jiězǔ*
       one   CLF:GENERAL   solution.group
       'a group of solutions'

    c. *yī*   *gè*   *yǒuqù*   *de*   *guāngmáng*
       one   CLF:GENERAL   interest   MOD   light.ray/trace
       'a trace of interest'

While kind classifiers typically occur with uncountable nouns (see Table 13 and Table 23), kind classifiers collocated with countable nouns and nouns annotated with [±count] are more likely to be equivalent to measure words in the translation between Chinese and English. As shown in Table 32, among the 98 measure words, 35 and 44 are equivalent to kind classifiers in the collocation with countable nouns and nouns annotated with

[±count], respectively. In contrast, only 19 of them are equivalent to kind classifiers collocated with uncountable nouns.

Table 32. The distribution of the direct translation of kind classifiers in Corpus 1

	direct translation of kind classifiers	
	#	%
+count	35	35.71
-count	19	19.39
±count	44	44.90
total	98	

Kind classifiers are more likely to be equivalent to general measure words when they are directly translated. Among the directly translated kind classifiers, they tend to be equivalent to *kind* (72 out of 98), and a few of them are also equivalent to such measure words as *type* (8), *sort* (2), *class* (5), and *group* (2), or to such more specific ones as *version* (1), *cohort* (1) and *variety* (1). As shown in example (85), *zhǒng* 'CLF:KIND, GENERAL' in (a) and (b) is comparable with *kind* and *type,* respectively, while *kuǎn* 'CLF:STYLE' in (c) is equivalent to *version* correlated with a commodity (watch), *lèi* 'CLF:SORT' in (d) is equivalent to *cohort* to refer to a group of people, and *zhǒng* 'CLF:KIND, GENERAL' in (e) is equivalent to *variety* pre-modified by the adjective *new* to show that the chocolate is new and different. However, in general, the equivalents of kind classifiers tend to be general without indicating the specific features of their referents.

(85) English equivalents of Chinese kind classifiers
    a. *yī*   *zhǒng*   *tiānrán*   *shíyòng*   *sèsù*
       one   CLF:KIND, GENERAL natural   edible   coloring
       'a kind of natural food colouring'
    b. *yī*   *zhǒng*   *wěi*   *jùlí*
       one   CLF:KIND, GENERAL pseudo-   distance
       'one type of pseudo-distance'
    c. *yī*   *kuǎn*   *méiyǒu*   *páizǐ*   *de*   *shǒubiǎo*
       one   CLF:STYLE   no   brand   mod   watch
       'an unbranded version of this(watch)'
    d. *yī*   *lèi*   *bìngrén*
       one   CLF:SORT   patient
       'cohorts of patients'
    e. *yī*   *zhǒng*   *nàirè*   *qiǎokèlì*
       one   CLF:KIND, GENERAL heat-proof   chocolate
       'a new variety of chocolate that withstands temperatures (up to 55℃)'

Very few event classifiers are shown as equivalence to measure words in English. Among the limited number of English measure words, they tend to indicate events, typically when

the nouns they cooccur with refer to entities rather than events. Example (86) shows the direct translation of event classifiers.

(86) The direct translations of event classifiers
    a. *yī*   *pán*   *qí*
          one   CLF:PLATE, GAME chess
          'a game of chess'
    b. *yī*   *lún*   *gāo'ěrfū*
          one   CLF:WHEEL, ROUND golf
          'a round of golf'
    c. *yī*   *zhèn*   *gǔshēng*
          one   CLF:PERIOD   drum.sound
          'a roll of drum'

As indicated in example (86), *pán* 'CLF:PLATE, GAME' in (a), *lún* 'CLF:WHEEL, ROUND' in (b), and *zhèn* 'CLF:PERIOD' in (c) are equivalent to *game*, *round* and *roll*, respectively, to indicate events, while the nouns they cooccur with refer to objects and sound respectively.

    To conclude, while numeral classifiers are usually said to be omitted when translated into non-classifier languages (See, e.g., Greenberg 1974: 84), a small proportion (3.04%) of Chinese numeral classifiers are shown as equivalence to measure words in English in Corpus 1 without the presence of adjectives. They are more likely to be specific entity classifiers and the general kind classifier. Furthermore, they are more likely to cooccur with countable nouns and have a quantity reading. Compared with the specificity of numeral classifiers in Chinese, their equivalent measure words in English tend to be more general in the context without adjectives, although a few numeral classifiers can also be equivalent to more specific measure words based on the features of noun referents.

6.3.1.2. In the context with adjectives in Corpus 2

In the context of adjectives, numeral classifiers are more likely to be shown as equivalence to measure words in English. Numeral classifiers pre-modified by adjectives tend to be shape classifiers, as discussed in §5.2, and are typically used with adjectives indicating size and shape, as discussed in §5.4.1. As shown in Table 33, among the 304 numeral classifiers equivalent to English measure words, 296 (97.37%) are entity classifiers

denoting shape, while the remaining eight include other specific entity classifiers (0.99%) and kind classifiers (1.64%).

Table 33. The distribution of direct translations of different types of Chinese numeral classifiers in Corpus 2

type of numeral classifiers		direct translation of numeral classifiers	
		#	%
entity	shape	296	97.37
	others	3	0.99
kind		5	1.64
total		304	

The entity classifiers directly translated in the context of adjectives are more likely to be collocated with countable nouns. Among the 375 entity classifiers, 203 (67.89%) are equivalent to measure words in English, as shown in Table 34. However, comparatively fewer entity classifiers collocated with uncountable nouns and nouns annotated as [±count] are equivalent to English measure words.

Table 34. The distribution of the direct translation of entity classifiers in Corpus 2

	direct translation of entity classifiers	
	#	%
+count	203	67.89
-count	47	15.72
±count	53	17.73
total	299	

The higher proportion of direct translation of entity classifiers with countable nouns can also be attributed to the unbalanced number of countable and uncountable nouns in Corpus 2. More importantly, entity classifiers collocated with an adjective are more likely to contribute new meanings to noun phrases (see §5.4.2). Thus, they are more comparable with measure words in English even when they are used with countable nouns. As shown in example (87), *píngguǒ* 'apple' and *huǒjī* 'turkey' collocated with the general classifier *gè* in (a) and (c) refer to an apple and a turkey, respectively. However, the entity classifiers *kuài* 'CLF:LUMP.LIKE' in (b) and *piàn* 'CLF:FLAT/THIN' in (d) used with the adjective *xiǎo* 'small, little' specify the shapes of their referents and refer to a small part of an apple and a piece of turkey (as food) respectively. While the general classifier *gè* can be omitted in the English translation without changing the meaning of the noun phrases in (a) and (c), the omission of *kuài* 'CLF:LUMP.LIKE' in (b) and *piàn* 'CLF:FLAT/THIN' in (d) can lead to different interpretations of the referents in English translation.

(87) Comparison of the semantic contribution of entity classifiers
  a. *yī*   *gè*   *píngguǒ*
     one   CLF:GENERAL   apple
     'an apple'
  b. *yī*   *xiǎo*   *kuài*   *píngguǒ*
     one   small, little CLF:LUMP.LIKE   apple
     'a bit of apple'
  c. *yī*   *gè*   *huǒjī*
     one   CLF:GENERAL   turkey
     'a turkey'

  d. *yī*   *xiǎo*   *piàn*   *huǒjī*
     one   small, little CLF:FLAT/THIN   turkey
     'a little piece of turkey'

In the context of adjectives, the direct equivalents of numeral classifiers are more likely to be measure words denoting more or less specific features of referents instead of only quantity. As shown in Table 29, only 5.74% of numeral classifiers in the context of adjectives are comparable with such equivalents as *bit, amount, much, few, volume* and *minimum*, which only denote quantity. However, 52.39% are equivalent to such measure words as *plot*, *lump, layer*, *block*, and *scrap* that indicate properties related to shape and size.

Numeral classifiers in the context of adjectives are most frequently equivalent to such general measure words as *piece* (98/274). Similar to numeral classifiers in Corpus 1, as mentioned above, *piece* is shown as equivalence to various entity classifiers, typically such shape classifiers as *kuài* 'CLF:LUMP.LIKE' (60/274) and *piàn* 'CLF:FLAT/THIN' (31/274), and collocated with a wide range of nouns. Furthermore, it does not reflect the specific shapes denoted by these numeral classifiers. As shown in example (88), all the three shape classifiers *kuài* 'CLF:LUMP.LIKE' in (a) and (c), *piàn* 'CLF:FLAT/THIN' in (b), and *tiáo* 'CLF: SLENDER' in (d) are equivalent to *piece*.

(88) Chinese numeral classifiers equivalent to the measure word *piece* in Corpus 2
  a. *yī*   *xiǎo*   *kuài*   *dàngāo*
     one   small, little CLF:LUMP.LIKE   cake
     'a small piece of cake'
  b. *yī*   *xiǎo*   *piàn*   *dàngāo*
     one   small, little CLF:FLAT/THIN   cake
     'a small piece of cake'
  c. *yī*   *xiǎo*   *kuài*   *mùtóu*
     one   small, little CLF:LUMP.LIKE   wood
     'a piece of wood'

d. *yī*   *cháng*   *tiáo*   *mùtóu*
   one   long   CLF:SLENDER   wood
   'a longer piece of wood'

Therefore, the specific properties denoted by the three shape classifiers are not reflected in the English translation, typically in examples (a) and (b), while the shape shown by *tiáo* 'CLF:SLENDER' is partly conveyed through the modifier *longer* in the English translation in (d), *longer* is more directly equivalent to the adjective *cháng* 'long' in Chinese than the numeral classifier *tiáo* 'CLF:SLENDER'. Other general measure words equivalent to entity classifiers include *portion* (2), *fraction* (1), *sort* (1) and *kind* (1). The only kind classifier used with adjectives, i.e., *lèi* 'CLF:CATEGORY', is equivalent to either *kind* in (a) or *class* in (b) to denote kind or type, as shown in example (89).

(89) The direct translations of *lèi*
   a. *yī*   *lèi*   *fēixiànxìng*   *piānwēifèn*   *fāngchéng*
      one   CLF:CATEGORY   nonlinear   partial.differential   equation
      'a class of nonlinear partial differential equations'
   b. *yī*   *lèi*   *fēilízǐxíng*   *biǎomiànhuóxìngjì*
      one   CLF:CATEGORY   non-ion   surfactant
      'a kind of non-ion surfactants'

On the other hand, some numeral classifiers in the context of adjectives can also be equivalent to more specific measure words. In Corpus 2, eleven entity classifiers are equivalent to 35 specific measure words in 163 numeral noun phrases. Take *kuài* 'CLF:LUMP.LIKE' as an example. It is comparable with 33 different measure words, aside from such general ones as *piece* and *bit*. As shown in example (90), *kuài* 'CLF:LUMP.LIKE' is equivalent to *lump* in English when used with such nouns as *niántǔ* 'clay' as in (a), but to *plot, parcel* and *patch* when used with nouns typically related to the land as in (b), *loaf* and *slice* when cooccurring with nouns related to bread as in (c), and *mouthful* and *morsel* when collocated with nouns related to food as in (d). Some of its English equivalents are also derived from its head nouns or implied features. For example, the equivalent of *kuài* 'CLF:LUMP.LIKE' in (e), i.e. *fragment*, is derived from the head noun *suìpiàn* 'fragment'. Similarly, *realm* in (f) is derived from the implied features related to work, with the noun *gōngzuò* 'work, job' shown as equivalence to more specific nouns *practice* and *management*.

(90) The translation of *kuài* 'CLF:LUMP.LIKE' in the context of adjectives

    a. *yī*   *xiǎo*   *kuài*   *níluóhé*   *niántǔ*

       one   small, little   CLF:LUMP.LIKE Nile.River   clay

       'a single lump of Nile River clay'

    b. *yī*   *xiǎo*   *kuài*   *dì*

       one   small, little   CLF:LUMP.LIKE *land*

       'a plot/patch/parcel/piece of land'

    c. *yī*   *xiǎo*   *kuài*   *xīnxiān de* *miànbāo*

       one   small, little   CLF:LUMP.LIKE fresh   MOD bread

       'a small loaf of fresh bread'

    d. *yī*   *xiǎo*   *kuài*   *shíwù*

       one   small, little   CLF:LUMP.LIKE food

       'a morsel of food'

    e. *yī*   *xiǎo*   *kuài*   *fúdòng huósāi suìpiàn*

       one   small, little   CLF:LUMP.LIKE floating piston *fragment*

       'a fragment of floating piston'

    f. *yī*   *xiǎo*   *kuài*   *gōngzuò*

       one   small, little   CLF:LUMP.LIKE work, job

       'this narrow realm of practice and management'

To conclude, compared with the translation of numeral classifiers in the context without adjectives, a significantly larger proportion of numeral classifiers (58.13%) in the context of adjectives are equivalent to English measure words. These English measure words tend to be more varied to express more specific features of their referents related to shape, size and other salient properties. Furthermore, measure words in English translation can also be equivalent to Chinese nouns instead of numeral classifiers in the context with adjectives.

6.3.2. Specific properties reflected in the English translation of Chinese nouns

While most numeral classifiers are omitted or equivalent to more general measure words in English typically in the context without adjectives, their specificity may be reflected in the translation of nouns. This section will examine the translation of nouns in numeral noun phrases and investigate how the English translation of nouns is used to express the specificity of referents denoted by numeral classifiers. The analysis will be made based on individuation hierarchy as well as the variation in the use of numeral classifiers Chinese as shown in §5.3.4 and §5.3.5.

Other than the general classifiers *gè* 'CLF:GENERAL', *zhī* 'CLF:SINGLE' and *zhǒng* 'CLF:KIND, GENERAL', specific classifiers are also assigned to different categories of nouns based on their semantic properties, as shown in Table 23. In the context without adjectives in Corpus 1, human and animal classifiers are used only with nouns for humans and animals, while other specific entity classifiers tend to be used with nouns with reference to bounded inanimates, as discussed in §5.3. In the context with adjectives in Corpus 2, specific entity classifiers tend to be used with nouns for bounded inanimates, as shown in §5.4. In this section, I will compare Chinese nouns with their English equivalents mainly in terms of four categories: nouns for humans, animals, bound inanimates, and neutralized entities.

The typical specific classifiers used with nouns for humans are *wèi* 'CLF:INDIVIDUAL, RESPECT' and *míng* 'CLF:IDENTITY', as shown in §5.3.3.1. Compared with the general classifier *gè*, *wèi* 'CLF:INDIVIDUAL, RESPECT' expresses higher social status of the referents, and *míng* 'CLF:IDENTITY' specifies the identities of the referents concerned. However, the English equivalents of human nouns in Chinese do not show such differences in terms of the specificity related to social status and identity. For example, the noun *zuòjiā* 'writer' is equivalent to *writer*, and *xuéshēng* 'student, pupil' is equivalent to *student* or *pupil*, no matter whether they are used with the general classifier *gè* or the specific classifiers *wèi* 'CLF:INDIVIDUAL, RESPECT' and *míng* 'CLF:IDENTITY'. Likewise, *cānyìyuán* 'senator' and *qǐgài* 'beggar' used with the above three numeral classifiers are all equivalent to *senator* and *beggar*, respectively.

While nouns for humans may have different equivalents in English, the different translations are largely based on the noun referents rather than the specificity shown in numeral classifiers. As shown in example (91), while *yīshēng* 'doctor, physician, surgeon' used with different numeral classifiers is equivalent to *doctor* in (a)-(c), it is also equivalent to *surgeon*, *physician* and *resident doctor*. However, the different choice of nouns in English are not determined by the specificity denoted by the numeral classifiers it cooccurs with, as *yīshēng* 'doctor, physician, surgeon' used with the general classifier *gè* and *wèi* 'CLF:INDIVIDUAL, RESPECT' in (d)-(g) can also be equivalent to the more generic noun *doctor* or more specific ones *surgeon*, *physician* or *resident* (doctor). Such translations are more likely to be determined by the roles of the referents indicated by the context, which will be further analysed in §7.3.

(91) The English translation of *yīshēng* 'doctor' with different human classifiers
a. *yī   gè*   *yīshēng*
   one CLF:GENERAL   doctor, physician, surgeon
   'A doctor'
b. *yī   míng*   *yīshēng*
   one CLF:IDENTITY   doctor, physician, surgeon
   'A doctor'
c. *yī   wèi*   *yīshēng*
   one CLF:GENERAL   doctor, physician, surgeon
   'One doctor'
d. *yī   wèi   wàikē*   *yīshēng*
   one CLF:GENERAL surgical   doctor, physician, surgeon
   'A surgeon'
e. *yī   wèi   jiéchū   de   yīshēng*
   one CLF:GENERAL distinguished MOD doctor, physician, surgeon
   'A distinguished physician'
f. *yī   gè   nénggàn   de   wàikē   yīshēng*
   one CLF:GENERAL able   MOD surgical doctor, physician, surgeon
   'an able surgeon'
g. *yī   gè   zhùyuàn*   *yīshēng*
   one CLF:GENERAL resident   doctor, physician, surgeon
   'A resident'

Only one non-human specific classifier, *shēng* 'CLF:SOUND', is used with human nouns in Corpus 1. As shown in example (92), the event classifier *shēng* 'CLF:SOUND' is used with human noun *bàba* 'dad, daddy, papa'. While the phrase can be equivalent to 'one call of daddy' in English translation provided in Corpus 1, the event classifier *shēng* 'CLF:SOUND' is omitted, while the noun *bàba* 'dad, daddy, papa' is equivalent to *daddy*. However, the Chinese noun *bàba* 'dad, daddy, papa' and its English equivalent *daddy* can be used as a form of address, and thus the two phrases in the translation between Chinese and English in Corpus 1 both refer to a call of the father concerned.

(92) The use of non-human classifiers with human nouns
   yī   shēng   *bàba*
   one   CLF:SOUND   dad, daddy, papa
   'One daddy'

Animal classifiers, either general or specific, tend to be omitted in English translations. Specific animal classifiers tend to denote some salient features or shape of their referents, as mentioned in §3.2.2. For example, *tóu* 'CLF:HEAD' and *wěi* 'CLF:TAIL' denote salient features of the referents, and *tiáo* 'CLF:SLENDER' specifies a slender shape of the referents. However, with the omission of animal classifiers in English translation, the specific

features they express are not reflected in English translations. For example, *dàxiàng* 'elephant' collocated with *zhī* 'CLF:SINGLE' and *tóu* 'CLF:HEAD' in both cases is equivalent to *elephant*, *gǒu* 'dog' used with *zhī* 'CLF:SINGLE' and *tiáo* 'CLF:SLENDER' is equivalent to *dog*, and *yú* 'fish' cooccurring with *tiáo* 'CLF:SLENDER' and *wěi* 'CLF:TAIL' is equivalent to *fish*.

Other specific classifiers used with animal nouns are event classifiers, and their specific features may be shown in the English equivalents of nouns. Only one event classifier, *dùn* 'CLF:SPELL, SESSION', is found to be used with animal nouns in Corpus 1. As shown in example (93), while *mǔlì* 'oyster' in (a) is equivalent to *oyster* in singular form when used with the general animal classifier *zhī* 'CLF:SINGLE', it is equivalent to *oysters* in plural form in (b), as it is used with the event classifier *dùn* 'CLF:SPELL, SESSION' with reference to a meal with oysters.

(93) The use of different numeral classifiers with *mǔlì* 'oyster'
    a. *yī*   *zhī*   *mǔlì*
       one   CLF:SINGLE   oyster
      'an oyster'
    b. *yī*   *dùn*   *mǔlì*
       one   CLF:SPELL, SESSION   oyster
      'oysters'

As to nouns for other bounded entities, specific features denoted by entity classifiers also tend to be omitted in their English translations both in the context with and without adjectives. Most nouns for bounded entities used with different numeral classifiers are their direct equivalents in English without showing the specificity denoted by numeral classifiers. For example, *shǒubì* 'arm' is equivalent to *arm* whether used with the two more general classifiers *gè* and *zhī* 'CLF:SINGLE' or the shape classifier *tiáo* 'CLF:SLENDER', and *zhàngpeng* 'tent' is equivalent to *tent* when used with either the general classifier *gè* or the specific classifier *dǐng* 'CLF:TOP'.

Similarly, specificity expressed by entity classifiers is not shown in the translation of most nouns for bounded concepts and events. As shown in example (94), *bàogào* 'report' is equivalent to *report* no matter whether it is used with the more general classifiers *gè* and *fèn* 'CLF:SHARE, PORTION', the more specific ones *xiàng* 'CLF:ITEM, PROJECT', *piān* 'CLF:ARTICLE' or the event classifier *cì* 'CLF:TIME'. Therefore, the specific features denoted by *xiàng* 'CLF:ITEM, PROJECT' indicating a project or a study in (c), *piān*

'CLF:ARTICLE' indicating a written report in (d), and *cì* 'CLF:TIME' indicating an event of giving a report in (e), are not shown in the English translations.

(94) The English translations of *bàogào* with different entity classifiers

a. *yī*   *gè*   *chūbù*   *de*   *bàogào*
one   CLF:GENERAL   preliminary MOD   report
'a preliminary report'

b. *yī*   *fèn*   *Éguó*   *de*   *bàogào*
one   'CLF:SHARE, PORTION Russia   MOD   report
'a Russian report'

c. *yī*   *xiàng*   *bàogào*
one   CLF:ITEM, PROJECT   report
'a report'

d. *yī*   *piān*   *yánjīu*   *bàogào*
one   CLF:ARTICLE   study   report
'a study'

e. *yī*   *cì*   *shūmiàn* *gōngjù jìnzhǎn* *bàogào*
one   CLF:TIME   written   tool   progress report
'a written tool progress report'

The specific meanings denoted by different entity classifiers tend to be omitted in their English translations even when they are used with the same nouns. For example, *méigui* 'rose' is equivalent to *rose* whether used with *duǒ* 'CLF:FLOWER.LIKE' or *zhī* 'CLF:BRANCH', although the former refers to a flower, while the latter indicates the branch where the flower is blooming. Similarly, *cāntīng* 'restaurant' cooccurring with such specific entity classifiers as *jiā* 'CLF:HOUSEHOLD' and *jiān* 'CLF:ROOM' is in both cases equivalent to *restaurant*, although the two specific classifiers may indicate that the restaurants concerned are of different sizes, with the former indicating a relatively larger restaurant. In the context of adjectives, *cǎodì* 'grassland, meadow' is equivalent to *grass* when it is used with both *piàn* 'CLF:FLAT/THIN' and *kuài* 'CLF:LUMP.LIKE', and *diànchí* 'battery' cooccurring with *kuài* 'CLF:LUMP.LIKE' and *lì* 'CLF:GRAIN.LIKE' is in both cases equivalent to *battery*, although the three numeral classifiers indicate different shapes. While the same nouns used with different numeral classifiers may be equivalent to different nouns in English, the choice of the nouns in English translation is more dependent on their modifiers rather than the specific features expressed by numeral classifiers. As indicated in example (95), *yǎnchū* 'show, performance' is equivalent to *show* in (a), *performance* in (b), *concert* in (c), and *tour* in (d). While *show* and *performance* can both be regarded as the equivalents of *yǎnchū* 'show, performance', *concert* in (c) and *tour* in (d) are not the

equivalents of *yǎnchū* 'show, performance' and the specificity expressed by the two nouns is derived from their modifiers rather than the two numeral classifiers.

(95) The English equivalents of *yǎnchū* with different entity classifiers
a. *yī*   *gè*   *kèrén*   *hùwài*   *yǎnchū*
one   CLF:GENERAL   *guest*   outdoor   show, performance
'an outdoor show'
b. *yī*   *cháng*   *fǎnfùwúcháng*   de   *yǎnchū*
one   CLF:VENUE   inconsistent   MOD   show, performance
'another inconsistent performance'
c. *yī*   *cháng*   *shìnèi*   *yuè*   *yǎnchū*
one   CLF:VENUE   chamber music   show, performance
'a chamber concert'
d. *yī*   *cì*   *xǐjù*   *xiànchǎng* *xúnhuí* *yǎnchū*
one   CLF:TIME   comedy live   tour   show, performance
'a live comedy tour'

Only a small number of nouns in English translation are found to express specific features denoted by numeral classifiers. For example, *chuán* 'ship, boat, vessel' can refer to a small boat, a ship, or vessel of larger size. It is more likely to be equivalent to *vessel* indicating a larger ship in English when it is used with the numeral classifier *sōu* 'CLF:SHIP', and to *ship* or *boat* when it cooccurs with *zhī* 'CLF:SINGLE' and *tiáo* 'CLF:SLEN-DER'. Another example concerned is the translation of *húzi* 'beard, moustache'. As shown in example (96), it can be equivalent to either *beard* or *moustache* as shown in (a)-(c). However, it is more likely to be equivalent to *moustache* when used with the numeral classifier *piě* 'CLF:LEFT.FALLING', typically when the adjective *xiǎo* 'small, little' is present either before or after it, as shown in (b) and (d). *Húzi* 'beard, moustache' is equivalent to more specific nouns or noun phrases as *pointed beard* in (e) when used with *liǔ* 'CLF:TUFT, LOCK' and into *soul patch* in (f) typically referring to the beard under the lower lip when used with *kuài* 'CLF:LUMP.LIKE'.

(96) The English translations of *húzi* with different entity classifiers
a. *yī*   *bǎ*   *húzi*
one   CLF:HANDLE   beard, moustache
'a moustache'
b. *yī*   *piě*   *xiǎo*   *húzi*
one   CLF:LEFT.FALLING small, little   beard, moustache
'a moustache'
c. *yī*   *fù*   *dà*   *húzi*
one   CLF:SET, ATTITUDE *big, large*   beard, moustache
'a big beard'

157

d. *yī*   *xiǎo*   *piě*   *húzi*
   one   small, little   CLF:LEFT.FALLING beard, moustache
   'that toothbrush moustache'

e. *yī*   *xiǎo*   *liǔ*   *húzi*
   one   small, little   CLF:TUFT, LOCK   beard, moustache
   'a little pointed beard'

f. *yī*   *xiǎo*   *kuài*   *húzi*
   one   small, little   CLF:LUMP.LIKE   beard, moustache
   'soul patch'

Nouns in English translation can also reflect the specific features shown by numeral classifiers and the referents of the noun phrase as a whole. As shown in example (97), *běn* 'CLF:BOOK' in (b) refers to a book and *jiā* 'CLF:HOUSEHOLD' in (e) refers to an organization. Therefore, the noun phrase in (b) is equivalent to *a notebook* instead of *a note* as shown in (a), and the noun phrase in (e) is equivalent to *a club* instead of *a team* as shown in (c) and (d).

(97) The translation of the specificity of referents expressed by numeral classifiers

   a. *yī* *gè*   *bǐjì*
     one CLF:GENERAL   *note*
     'a note'

   b. *yī* *běn*   *bǐjì*
     one CLF:BOOK   note
     'a notebook'

   c. *yī* *gè*   *qíuduì*
     one CLF:GENERAL   team
     'a team'

   d. *yī* *zhī*   *hénqiáng*   *de*   *qíuduì*
     one CLF:BRANCH   very.strong   MOD   team
     'a very strong team'

   e. *yī* *jiā*   *zhēnzhèng* *xūyào* *wǒ* *qù* *de*   *qíuduì*
     one CLF:HOUSEHOLD really   need me go MOD team
     'a club who wants me'

Compared with other groups of nouns, polysemous nouns in Chinese are more likely to be reflected by specific nouns in English translation, as their referents can be differentiated by numeral classifiers based on the specific features. As indicated in example (98), *biānjí* 'edit, editor' is equivalent to *writer* in (a) based on the human classifier *míng* 'CLF:IDENTITY', *juésè* 'character, role, part' is equivalent to 'character' in (b) based on the human classifier *wèi* 'CLF:INDIVIDUAL, RESPECT', and *zúzhī* 'organization, tissue' is equivalent to *group* and *tissue* based on the two entity classifiers *jiā* 'CLF:HOUSEHOLD' in (c) and *tiáo* 'CLF:SLENDER' in (d), respectively.

(98) The translation of polysemous nouns

    a. *yī*   *míng*   *biānjí*
        one CLF:IDENTITY   edit, editor
        'a writer'

    b. *yī*   *wèi*   *yǎnshēng* *yǔzhòu*   *juésè*
        one CLF:INDIVIDUAL, RESPECT expanded   universe   character, role, part
        'an expanded universe character'

    c. *yī*   *jiā*   *guójì*   *gāngtiěyè* *zúzhī*
        one CLF:HOUSEHOLD   international steel.industry organization, tissue
        'a group'

    d. *yī*   *tiáo*   *zúzhī*
        one CLF:SLENDER   organization, tissue
        'a band of tissue'

To conclude, specific properties denoted by numeral classifiers are not very frequently reflected in the choice of nouns in English, even when numeral classifiers are omitted in the translation. Most nouns for humans, animals and inanimates are comparable with the same equivalents in English when they are used with different numeral classifiers. The choice of different nouns in English translation mainly occur in three situations: first, specific properties of the referents are denoted by nouns or the modifiers of nouns instead of numeral classifiers; second, numeral classifiers contribute new meanings to noun phrases and thus the referents should be interpreted based on all the elements rather than one single element in the phrase; third, nouns are polysemous or generic and their referents should be differentiated based on the specific properties denoted by numeral classifiers.

6.4. Ascribing properties to referents

The function concerning ascribing properties to referents is related to social status and affective meanings expressed by numeral classifiers, as shown in §5.3.5. The variation in the use of classifiers to convey such meanings can be interpreted in terms of reassignment of classifiers to referents ranked on different levels on the individuation hierarchy (see Table 23). This section will examine how this function is represented in the translation in Corpus 1. General classifiers related to human, animal, bounded inanimates and unbounded entities will be examined first, followed by the more specific classifiers on the order of the individuation hierarchy.

Three general classifiers are ordered on the individuation hierarchy: *gè* 'CLF:GENERAL', *zhī* 'CLF:SINGLE' , and *zhǒng* 'CLF: KIND, GENERAL'. As shown in §5.3.4, *gè* is in opposition to *zhī* by ranking on the top of the individuation hierarchy for human referents, while *zhī* is ranked as the second on the hierarchy with reference to animals. As to the opposition between *gè* and *zhǒng* when used for inanimates, the former is the general classifier for bounded entities, while the latter is the general classifier for mass entities. The switch of the general classifiers may upgrade or downgrade the referents on the individuation hierarchy and thus express messages related to social status and affective meanings, as shown in §3.3.1.

The classifier *zhī* 'CLF:SINGLE' does not cooccur with human nouns in Corpus 1, although among native speakers, it can be used with reference to kids or youngsters. For example, the band TFBoys consisting of three boys tends to be referred to by their fans in China as *sān xiǎo zhī* (three little CLF:SINGLE) 'three little boys'. While using the general animal classifier with nouns for humans downgrades the referent on the individuation hierarchy, here its use can be interpreted as an expression of affection. Since no related example is found in Corpus 1, it can be assumed that such properties tend to be omitted when translated into English.

In Corpus 1, a small proportion (6.81%) of nouns for animals are used with the general classifier *gè* when their referents are either personified or more distinguishable, which can be interpreted in terms of upgrading on the individuation hierarchy (see §5.3.5). However, such properties are not shown in the English translation. As shown in example (99), *gǒu* 'dog' is equivalent to *dog* when used with the general animal classifier *zhī* 'CLF:SINGLE' in (a), the more specific shape classifier *tiáo* 'CLF:SLENDER' in (b), or the general classifier *gè* in (c). While *xiǎogǒu* 'little.dog' in (c) is personified by being used with the general classifier *gè*, the ascribed property is not shown in the English translation. Similarly, *dàxiàng* 'elephant' is equivalent to *elephant* whether it is used with the two animal classifiers *zhī* 'CLF:SINGLE' in (d) and *tóu* 'CLF:HEAD' in (e) or the general classifier *gè* in (f).

(99) Comparison of the English equivalents of nouns for animals with animal classifiers and the general classifier *gè*
    a. *yī*   *zhī*   *gǒu*
       one   CLF:SINGLE   dog
       'a dog'

b. *yī*   *tiáo*   *xiǎo*   *tǐxíng*   *de*   *gǒu*
one   CLF:SLENDER   small, little   figure   MOD   dog
'a small dog'

c. *yī*   *gè*   *bènbèn*   *de*   *máorōngrōng*   *de*   *xiǎogǒu*
one   CLF:GENERAL   silly   MOD   fluffy   MOD   little.dog
'a little soft fluffy dog'

d. *yī*   *zhī*   *chéngnián*   *Yàzhōu*   *dàxiàng*
one   CLF:SINGLE   full-grown   Asian   elephant
'a full-grown Asian elephant'

e. *yī*   *tóu*   *dàxiàng*
one   CLF:HEAD   elephant
'an elephant'

f. *yī*   *gè*   *xīn*   *bǔhuò*   *de*   *dàxiàng*
one   CLF:GENERAL   new   capture   MOD   elephant
'a newly captured elephant'

While the use of *zhǒng* with countable concrete nouns may indicate downgrading of the referents as unspecified mass on the individuation hierarchy (see §5.3.5), it may involve derogation of the referents when used with human nouns. The only one example with the collocation of the general kind classifier with human noun in Corpus 1 is shown in example (100). By being used with the general classifier, the referent of the noun *zhǔfù* 'housewife' is less individuated and the noun is equivalent to *homebody* in English translation, a more disparaging noun compared with *housewife*.

(100) Translation of Chinese countable concrete nouns collocated with *zhǒng*
yī   *zhǒng*   *bùzhébùkòu*   *de*   *jiātíng*   *zhǔfù*
one   CLF:KIND, GENERAL   complete   MOD   family   housewife
'A complete homebody'

The typical specific classifiers used to ascribe properties to referents are *wèi* 'CLF:INDIVIDUAL, RESPECT' and *zūn* 'CLF:RESPECT' in opposition to the general classifier *gè*, *zhī* 'CLF:SINGLE', or other specific classifiers. The classifiers *wèi* 'CLF:INDIVIDUAL, RESPECT' and *gè* can show a different social status and express affective meanings such as respect and contempt. In contrast, *zūn* 'CLF:RESPECT' can ascribe additional value to referents. These features are rarely reflected in the English translations in Corpus 1. As shown in example (101), *qǐgài* 'beggar' in (a) and (b) is equivalent to *beggar*, while *guānyuán* 'official' in (c) and (d) is equivalent to *official*, with the upgraded status and respect attributed by *wèi* 'CLF:INDIVIDUAL, RESPECT' in (b) and (d) omitted in the translation.

(101) The English translation of human nouns used with *gè* and *wèi*
    a. *yī*   *gè*   *lăo*   *qĭgài*
       one   CLF:GENERAL   old   beggar
       'an old beggar'
    b. *yī*   *wèi*   *liùshíwŭ*   *suì*   *de*   *qĭgài*
       one   CLF:INDIVIDUAL, RESPECT   sixty-five   age   MOD   beggar
       'a 65-year-old beggar'
    c. *yī*   *gè*   *ăidūn*   *de*   *dāngdì*   *zhèngfŭ*   *guānyuán*
       one   CLF:GENERAL   pudgy   MOD   local   government   official
       'a pudgy local government official'
    d. *yī*   *wèi*   *měiguó*   *guānyuán*
       one   CLF:INDIVIDUAL, RESPECT   America   official
       'A US government official'

Similarly, *diāoxiàng* 'statue' collocated with *zūn* 'CLF:RESPECT' in (a) and *zuò* 'CLF:SEAT, PEDESTAL, BASE' in (b) in example (102) are both equivalent to *statue* without showing the additional value expressed by *zūn*. However, the noun *thing* in (c) can be regarded as a reflection of the degrading feature denoted by *gè* in opposition to *zūn* 'CLF:RESPECT'. It seems that the statue is not valuable enough to be collocated with *zūn* 'CLF:RESPECT' in Chinese or referred to as a statue in English.

(102) The English translation of nouns used with *zūn, zuò and gè*
    a. *yī*   *zūn*   *diāoxiàng*
       one   CLF:GENERAL   statue
       'a statue'
    b. *yī*   *zuò*   *jùdà*   *de*   *diāoxiàng*
       one   CLF:SEAT, PEDESTAL, BASE   giant   MOD   statue
       'a giant statue'
    c. *yī*   *gè*   *nísù*   *de*   *diāoxiàng*
       one   CLF:GENERAL   clay.carve   MOD   statue
       'a thing made of clay'

In conclusion, the function of Chinese numeral classifiers concerning ascribing properties to referents is rarely reflected in their English translations in Corpus 1. Most numeral classifiers are omitted in the translation and most nouns are reflected in their equivalents in English, with only a few exceptions occurring in the choice of nouns involving affective meanings related to social status or value. Therefore, the properties that numeral classifiers attribute to their referents are usually omitted in the translation.

6.5. Discussion and concluding remarks

This chapter has compared Chinese numeral classifiers and their related elements with their equivalents in English based on numeral noun phrases with and without adjectives. This section will give a brief summary of the findings regarding the direct equivalents of Chinese numeral classifiers, the representation of Chinese numeral classifiers in terms of their lexical and grammatical meanings, and finally, their semantic functions in English translation.

The results show that Chinese numeral classifiers are not always omitted in English translation. While only a small proportion (3.06%) of numeral classifiers are equivalent to English measure words in the context without adjectives, over a half of them (58.13%) are comparable with English measure words in the context of adjectives. In the context without adjectives, most numeral classifiers equivalent to English measure words are specific classifiers and kind classifiers, while the general classifier *gè* is rarely equivalent to measure words. This finding to a certain extent challenges the traditional description of numeral classifiers as omitted when translated into a non-classifier language (Greenberg 1978: 84). On the other hand, the results also confirm the hypotheses that the general classifier *gè* is more likely to be omitted in English translation, whereas specific classifiers are more frequently to be equivalent to English measure words, typically in the context with adjectives.

The grammatical features of numeral classifiers concerning definiteness are mainly shown in the translation of the numeral *yī* 'one'. The numeral *yī* 'one' is rarely correspondent to its equivalent *one* in English. Instead, over 80% of its occurrences are correspondent to indefinite articles and determiners to express indefiniteness and about 5% to the definite article, demonstratives and possessive structures to express definiteness.

The lexical meanings of numeral classifiers concerning specific properties can be reflected in the equivalent measure words or a small number of nouns in English translation. Direct equivalence of numeral classifiers with measure words occurs more frequently in the context with adjectives, as shown above. Numeral classifiers are also likely to be equivalent to measure words in English translation when they occur with nouns ranked either higher or lower on the individuation hierarchy. For example, entity classifiers used with uncountable nouns and kind classifiers cooccurring with countable nouns are more likely to be equivalent to measure words in English. Finally, numeral classifiers

can be equivalent to measure words in English when they contribute new meanings related to quantity to noun phrases. For example, the numeral classifier *céng* 'CLF:LAYER' is more likely to equivalent to measure words as it denotes the physical feature related to two dimensions of the referents and it also implies a quantity reading.

The semantic functions of numeral classifiers are reflected in English translation in several ways. The function concerning individuation of nouns is shown in English measure words or the translation of *yī* 'one'. In English translation, measure words are required for uncountable nouns, or when they are equivalent to numeral classifiers which contribute new meanings related to quantity. In terms of the translation of *yī* 'one', it is most frequently correspondent to articles and determiners and most nouns are shown in singular form in English translation. The indefinite reading of *yī* 'one' can sometimes be shown in the plural form of English nouns, when it is equivalent to the English determiner *every* or *each*, which appears in a singular form but has a plural reference. While countable nouns can be directly enumerated and thus an additional element, e.g., measure word, is not required, the plural suffix should be added to nouns when they have a plural reading in English. The results give support to the argument by Borer (2005: 94) and Her (2012: 29) that the plural suffix functions like numeral classifiers to individuate nouns, only that it can be omitted when the quantity is one.

The function involving differentiating referents is rarely reflected in either measure words or nouns in English. In the context without adjectives in Corpus 1, only 3.06% numeral classifiers are equivalent to such measure words as *piece*, *kind*, and *bit*, which are less specific compared with specific classifiers. While far more numeral classifiers (58.13%%) in the context with adjectives are equivalent to more specific measure words in English, numeral classifiers, except for shape classifiers, rarely occur with adjectives. The specificity expressed by numeral classifiers is seldom reflected in the choice of nouns in English, either. Most nouns are reflected in their equivalents in English when they are used with different numeral classifiers. Those equivalent to more specific ones in English tend to be polysemous or unspecified nouns in Chinese. These findings contradict the hypothesis that specific features denoted by numeral classifiers tend to be expressed by nouns in English if they are not directly translated into measure words. The results may be explained by that fact that most nouns in the numeral noun phrases in the two corpora are specific enough to entail the features expressed by the numeral classifiers.

Similarly, the function involving ascribing properties to referents is rarely shown in the English translation. The properties attributed to noun referents by numeral classifiers tend to be omitted in English translation. For example, with the omission of numeral classifiers in translation, nouns for animals are shown in their equivalents in English, whether they are used with the general classifier *gè* for human or inanimate referents or the general animal classifier *zhī* 'CLF:SINGLE', while human nouns are represented by their equivalents whether they are used with the general classifier *gè* or the specific classifier *wèi* 'CLF:INDIVIDUAL, RESPECT' involving affective meanings related to respect. This function is only represented in a few instances of different choice of nouns in English. Instead of being shown in their direct equivalents, Chinese nouns may be reflected in more general ones or nouns more derogatory in English to show the contempt or irony expressed by numeral classifiers.

In summary, Chinese numeral classifiers can be directly translated into English measure words typically when numeral classifiers are pre-modified by adjectives. However, the semantic functions of numeral classifiers are rarely represented in English translation. The function involving individuation of nouns is reflected in a limited number of measure words or plural nouns, other than the articles used with nouns in singular form with the omission of the plural suffix -*s* in English, according to Her (2012). With regard to differentiating referents, the specific properties are reflected by English measure words when their equivalent numeral classifiers occur with adjectives or by English nouns when their equivalent nouns in Chinese are polysemous or unspecified. Concerning the function of ascribing properties, it is only represented in a very limited number of instances of a different choice of nouns in English other than the direct equivalents of Chinese nouns.

Chapter 7: Chinese numeral classifiers in translation: discourse functions

7.1. Introduction

Numeral classifiers, like all the other types of nominal classification devices, play an important role in the organization of discourse, as mentioned in §2.4.2. For example, their presence or choice can be used to establish and manage the status of nominal referents. This chapter will examine how discourse functions of numeral classifiers are reflected in English translation. To be more specific, this chapter aims to address the following research questions based on Corpus 3: a) how are referents identified, managed, and recategorized in discourse in the data in Corpus 3?; and b) what equivalent forms in English are used to represent the functions involving reference identification, reference management, and representation of referents in discourse?

In order to address the issues, in §7.2 I will first briefly review the frequency of different types of numeral classifiers in Corpus 3. In §7.3, I will discuss the function concerning reference identification. Uses of numeral classifiers related to definiteness, referentiality and topicality will be examined in §7.4 and the function concerning recategorization of referents and its representation in English will be investigated in §7.5. Discussion and conclusions will be given in §7.6.

7.2. Frequency of different types of numeral classifiers in Corpus 3

This section will give a general introduction of the data in Corpus 3. Corpus 3 was compiled based on 411 pairs of sentences derived from five chapters of the novel *The Three-Body Problem* by Liu Cixin*,* the first Asian work of science fiction ever to win the Hugo Award for Best Novel in 2015. The five chapters chosen for the study serve as the introduction in the novel with the debut of some of the main characters, including Wang Miao, a nanomaterials scientist who finds the solution to the three-body problem, and Shi Qiang (nicknamed Da Shi), a controversial police detective and counter-terrorism specialist who is credited with the prevention of a terrorist bombing, and the re-introduction of one of the main characters, Ye Wenjie, an astrophysicist as well as the first person who contacts the Trisolarans and puts the Earth under the threat of their invasion.

In the corpus, there are 645 pairs of noun phrases, including 618 phrases involving the use of numeral classifiers and 27 numeral noun phrases without numeral classifiers. As shown in Table 35, numeral classifiers occur predominantly in the two constructions of [NUM/DEM/PRON+CLF+N] (80.62%) and [NUM/DEM/PRON+CLF] (10.08%). In these structures, the numeral *yī* 'one' can be omitted after the demonstratives or interrogative pronouns.

Table 35. The frequency of different types of numeral noun or classifier phrases in Corpus 3

structure	occurrence	
	#	%
[NUM/DEM/PRON+CLF+N]	520	80.62
[NUM/DEM+CLF]	65	10.08
[NUM/DEM/DET+N]	27	4.19
[CLF+N]	23	3.57
[NUM+CLF+CLF+(N)]	10	1.55
total	645	

Classifiers also occur in the construction of [CLF+N] (3.57%) and [NUM+CLF+CLF+(N)] (1.55%). The construction [CLF+N] usually immediately follows a verb, and the general classifier *gè* is a more frequent choice in the structure. As shown in example (103), the general classifier *gè* is used with the noun *tiāncái* 'genius' after the copula *shì* 'be' in (a), while *jiàn* 'CLF:PIECE' occurs with the noun phrase *zāngxī、ixi de píjiākè* (dirty MOD leather.jacket) 'dirty leather jacket' after the verb *chuān zhe* (wear PROG) 'be wearing' in (b). As to the construction [NUM+CLF+CLF+(N)] in

167

example (c), the numeral classifier *zuò* 'CLF:SEAT, PEDESTAL, BASE' is repeated to denote a plural meaning of *dānbǎi* 'pendulum'.

(103) The structures of [CLF+N] and [NUM+CLF+CLF(+N)]

    a. *tā*   *kě*   *shì*   **gè**   **tiāncái**
       he   exactly   be   CLF:GENERAL   genius
       'He's **a genius**'

    b. *chuān zhe*   **jiàn**   **zāngxixi**   *de*   **píjiākè**
       wear   PROG   CLF:PIECE   dirty   MOD   leather.jacket
       'wearing **a dirty leather jacket**'

    c. *yī*   *zuò*   *zuò*   *jùdà*   *de*   *dānbǎi*
       one   CLF:SEAT, PED-CLF:SEAT, PED-giant   MOD   pendulum
              ESTAL, BASE   ESTAL, BASE
       'the numerous giant pendulums'

Numeral classifiers may be omitted in a small number of noun phrases, as in the construction of [NUM/DEM+N] (4.19%). As shown in example (104), such noun phrases tend to occur as the subjects, e.g., *zhè rén* 'this man' in (a) or follow immediately the name of a person, e.g., *yī rén* 'one person' after the name *Wang Miao* in (b). They tend to refer to salient referents and occur very close to the related nouns or names concerned. Their discourse function involving thematic salience will be further analysed in §7.4.2.

(104) The structures of [NUM/DEM+N]

    a. ***Zhè***   ***rén***   *zěnme*   *zhèyàng?*
       this   man   how   in.this.way
       'What's wrong with **that guy**?'

    b. *bùzhǐ*   *shì*   *ràng*   *Wāng Miǎo*   **yī**   **rén**   *búshì*。
       not.limited be   let   Wang Miao   one   person   discomfort
       'Wang wasn't **the only one** annoyed by his rough manner.'

Corpus 3 and Corpus 1 show a similar distribution of types of numeral classifiers. Table 36 shows the frequency of different types of numeral classifiers in Corpus 3. While specific entity classifiers (44.34%) occur as the most frequent type of numeral classifiers, the two general classifiers *gè* (42.55%) and *zhǒng* 'CLF:KIND, GENERAL' (8.25%) are the two most frequently used numeral classifiers in the corpus. In comparison with Corpus 1 (see Table 10), specific entity classifiers occur more frequently in Corpus 3 (44.34% vs. 33.70%), which attests to the finding that specific classifiers occur more frequently in fiction according to Xiao and McEnery (2010: 49) and Erbaugh (1986: 403).

Table 36. The frequency of different types of Chinese numeral classifiers in Corpus 3

type of numeral classifiers		occurrence	
		#	%
entity	general: *gè*	263	42.55
	specific	274	44.34
event		26	4.21
kind	general: *zhǒng*	51	8.25
	specific	4	0.65
total		618	

To conclude, this section has given a brief introduction of the data in Corpus 3, including the types of numeral and classifier phrases and the frequency of different types of numeral classifiers. The types of numeral classifiers in Corpus 3 show a similar distribution to that in Corpus 1, except for a relatively higher percentage of specific entity classifiers in Corpus 3.

7.3. Reference identification

The discourse function of numeral classifiers concerning reference identification is related to anaphora, deixis and disambiguation. A classifier used anaphorically refers to a previously mentioned referent, while a classifier used deictically identifies a referent which has not been mentioned in the discourse. Numeral classifiers can also be used to disambiguate referents based on their specific features. This section will deal with the three functions and their representation in English translation based on Corpus 3.

Numeral classifiers are frequently used anaphorically to refer to previously mentioned entities. In Corpus 3, numeral classifiers used anaphorically constitute about 8.06% of all the occurrences and the general classifier *gè* (2.02%) accounts for only a quarter of the occurrences. Numeral classifiers for anaphoric use tend to occur in the construction of [NUM+CLF] without the presence of head nouns. Anaphoric use of the general classifier is illustrated in example (105).

(105) Anaphoric use of the general classifier in Corpus 3

a. ***Yǔzhòu***   ***bèijǐng***   ***fúshè,***   *nǐ*   *zěnme*
   Cosmic   background   radiation   you   how
   duì   ***zhè***   ***gè***   *yǒu*   *xìngqù?*
   to   this   CLF:GENERAL   have   interest
   'The cosmic microwave background? What made you interested in **that**?'

b. ...	*yǒu*	***sān***	***gè***	***ángguì***
…	have	three	CLF:GENERAL	expensive
de	***'táiqiúzhuō'***	*bèi*	*zào*	*le*
MOD	pool.table	BEI	construct	PRT
chūlái,	***yī***	***gè***	*zài*	*Běiměi,*
out	one	CLF:GENERAL	at	North.America
yī	***gè***	*zài*	*Ōuzhōu,*	*háiyǒu*
one	CLF:GENERAL	at	Europe	also
yī	***gè***	*nǐ*	*dāngrán*	*zhīdào,*
one	CLF:GENERAL	you	certainly	know
zài	*Zhōngguó*	*Liángxiāng*	…	
at	China	Liangxiang		

'…Three expensive '**pool tables**' have been constructed: **one** in North America, **another** in Europe, and **the third** you are familiar with, in Liangxiang…'

As shown in example (105), the general classifier *gè* is used to anaphorically refer to the two referents in subsequent mentions. In English translation, with the omission of the general classifier, referents are mentioned again by such pronouns *that* in (a) and *one*, as shown by *one*, *another* and *the third* in (b).

Specific classifiers are more likely to be used anaphorically. Among the 52 examples of such anaphoric use, there are 39 occurrences of specific classifiers. Examples are given in (106).

(106) Anaphoric use of specific classifiers without head nouns in Corpus 3

a. *Shǐqiáng*	*bǎ*	*shǒuzhōng*	*de*	*yāntóu*
Shiqiang	BA	in.hand	MOD	butt
rēngdiào,	*cóng*	*yābiǎn*	*de*	*yānhé*
throw,	from	flatten	MOD	cigarette.pack
lǐ	*chōuchū*	***yī***	***gēn。***	
in	draw	one	CLF:ROOT, STICK-SHAPE	

'Shi threw away the butt and took out **another cigarette** from a flattened pack.'

b. *Jìdé*	*zài*	*dàsān*	*de*	*yī*
remember	in	college.third-year	MOD	one
cì	*xìnxī*	*kè*	*zhōng,*	*jiāoshòu*
CLF:TIME	information	class	in	professor
guàchū	*le*	***liǎng***	***fú***	
put.up	PRT	two	CLF:PICTURE	
dà	***túpiàn,***	***yī***	***fú***	*shì*
big	picture	one	CLF:PICTURE	be
huàmiàn	*pángzá*	*jīngxì*	*de*	*Qīngmíng*
picture	rich	fine	MOD	Qingming
Shàng	*Hé*	*Tú,*	***lìng***	***yī***
on	river	painting	other	one

170

fú	*shì*	*yī*		*zhāng*
CLF:PICTURE	be	one		CLF:SPRED.OPEN/FLAT
kōngkuàng	*de*	*tiānkōng*	*zhàopiàn*	...
open	MOD	sky	picture	

'He remembered taking a class in information theory as a third-year student in college. The professor had put up **two pictures**: **One** was the famous Song Dynasty painting *Along the River During the Qingming Festival*, full of fine, rich details; **the other** was a photograph of the sky on a sunny day…'

In example (106), the numeral classifier *gēn* 'CLF:ROOT, STICK-SHAPE' in (a) is used to refer to the cigarette, even though the noun *yān* 'cigarette' is not mentioned , while in (b), there are two pictures, referred to again in presentative structures by the numeral classifier *fú* 'CLF:WIDTH PICTURE' one after the other. What should be noted is that numeral classifiers in the construction [NUM+CLF] used anaphorically are less likely to be replaced by other classifiers, including the general classifier *gè*. As shown in (b), the numeral classifier *fú* 'CLF:WIDTH PICTURE' is used consistently in the three phrases to track the referents. In English translation, the anaphoric use of numeral classifiers in the construction of [NUM+CLF] is more likely to be represented by more specific nouns as shown by *cigarette* in (a) or by such pronouns as *one* and *the other* as shown in (b).

The above examples have illustrated the anaphoric use of numeral classifiers to identify previously mentioned referents. A related way of the anaphoric use of numeral classifiers is to help to disambiguate between antecedents. As at least two referents need to be disambiguated, there should be at least two different numeral classifiers used to make the distinctions. Only one example related to the use for disambiguation is found in Corpus 3, which is provided in (107).

(107) Use of numeral classifiers to disambiguate between antecedents in Corpus 3

Chōng	*chūlái*	*de*	*jiāopiàn*	*shàng,*
develop	out	MOD	negative	on
nà	*shùzì*	*yōulíng*	*bān*	*de*
that	number	ghost	like	MOD
zài	**měiyī**	**zhāng**		**dǐpiàn**
at	each	CLF:SPREADING OPEN/FLAT		negative
shàng	*búduàn*	*xiǎnshì*	*chūlái,*	**dìyī**
on	continually	appear	out	first
zhāng		*shì*	*1187:27:39,*	*cóng*
CLF:SPREADING OPEN/FLAT		be	1187:27:39,	from
shàngyī	**juǎn**	**zuìhòu**	**yī**	**zhāng**

171

previous	CLF:ROLL	last	one	CLF:SPREADING OPEN/FLAT
pāishè	*dào*	*pāi*	***zhè***	***juǎn***
shoot	to	shoot	this	CLF:ROLL
de	***dìyī***	***zhāng*,**		*zhènghǎo*
MOD	first	CLF:SPREADING OPEN/FLAT		just
shì	*jiāngé*	*zhème*	*cháng*	*shíjiān。*
be	gap	so	long	time
Yǐhòu	*de*	***měiyī***	***zhāng***	
hearafter	MOD	each	CLF:SPRED.OPEN/FLAT	
de	*jìshí*	*jiāngé*	*wéi*	*sān*
MOD	timing	gap	be	three
dào	*sì*	*miǎo*	…	
to	four	second		

'In the developed **roll**, the numbers again appeared on every **negative** like ghosts. **The first one** was marked 1187:27:39. The difference matched the passage of time between **the last shot of the last roll** and **the first shot of this roll**. After that, the number decreased by three or four seconds in **each image**…'

As shown in example (107), there are two referents mentioned in the example: negative rolls and negatives, and they are referred to in the following mentions and disambiguated consistently by the two specific classifiers: *zhāng* 'CLF:SPREADING OPEN/FLAT' for the negatives and *juàn* 'CLF:ROLL' for the negative rolls. The results show that numeral classifiers used to disambiguate among referents are specific classifiers. Otherwise, ambiguity cannot be avoided if two referents are both referred to by the general classifier when head nouns are omitted. Furthermore, specific classifiers used to disambiguate referents should be consistent so that it makes possible to track the referents, and therefore, they are not likely to be replaced by the general classifier in subsequent mentions. In English translation, the function related to disambiguation is reflected in the choice of nouns. The negative rolls are referred to consistently by the noun *roll* in English. However, the negatives are referred to again in different ways by the pronoun *one* only for anaphoric use and the nouns *shot* and *image* to avoid ambiguity with the negative rolls.

Deictic use of numeral classifiers occurs in conversations to indicate referents that are obvious from context. There are three examples of deictic use of numeral classifiers in Corpus 3 and all of them occur in conversations. Among the three examples, one involves the use of a specific classifier and the other two the general classifier *gè*. Furthermore, all the numeral classifiers in the three examples occur with the demonstrative *zhè*

'this', while nouns are not explicitly mentioned. Two of the examples involving the use of the specific and general classifiers are illustrated in example (108).

(108) The deictic use of numeral classifiers in Corpus 3

a. …   *tā*   *gěi*   *Láikǎ*   *zhuāngshàng*
   …   he   give   Leica   load
   dìsān   *gè*   *jiāojuǎn,*   *bǎ*   *xiàngjī*
   third   CLF:GENERAL   film   BA   camera
   dìgěi   *qīzǐ:*   *"lái,*   *pāiwán*   **zhè**   **juǎn**。
   hand   wife   come   shot.finish   this   CLF:ROLL
   '…he loaded another roll of film in the Leica and handed it to his wife. "Here, finish **the roll** for me."'

b. *Wāng Miǎo*   *bǎ*   *Kēdá*   *xiàngjī*   *sāidào*
   Wang Miao   BA   Kodak   camera   stuff
   guòlái   *chīfàn*   *de*   *liù*   *suì*
   come   take.meal   MOD   six   year
   érzi   *shǒulǐ,*   *"Dòu Dòu,*   *nǐ*   *bāng*   *bàba*
   son   hand   Dou Dou   you   help   Daddy
   pāi。   *Jiù*   *àn*   **zhè**   **gè,**   *duì,*
   shot   just   push   this   CLF:GENERAL   right
   zhè   *shì*   *yī*   *zhāng*   …"
   this   be   one   CLF:SPREADING OPEN/FLAT
   'Wang stuffed the Kodak into the hands of his six-year-old son, who was about to start eating dinner. "Dou Dou, come help Daddy. Push **this button**. Right, like that. That's one shot…"'

As shown in example (108), the specific classifier *juǎn* 'CLF:ROLL' in (a) is used to identify the referent, i.e., a roll of films, when Wang Miao hands the camera to his wife, even though the referent has not been previously mentioned in the conversation. In contrast, in (b) the button on the camera is indexed by the general classifier *gè*, when Wang Miao points at the button and teaches his son how to handle the camera. In English translation, the deictic use of numeral classifiers is represented by the choice of specific nouns, as shown by *roll* in (a) and *button* in (b).

The results show that numeral classifiers, including both general classifiers and specific classifiers, can be used for reference identification. Specific classifiers occur more frequently in the construction without head nouns to anaphorically refer to their antecedents or to disambiguate referents, in which case they tend to be used consistently to avoid ambiguity. Both general and specific classifiers can be used deictically in conversations and they are more likely to occur with demonstratives, typically *zhè* 'this', to

identify a referent that does not occur in previous mentions. In English translation, these functions of numeral classifiers are more likely to be shown in the choice of nouns or pronouns. Specific nouns are more frequently used for deixis and disambiguation, while such pronouns as *that* and *one* and *the other* are more likely to occur for anaphora.

7.4. Reference management

The discourse function of numeral classifiers discussed here involves reference management, which is related to definiteness, referentiality and topicality, as discussed in §3.3.2. This section will examine how numeral classifiers are used to indicate definiteness and referentiality and exhibit discourse patterns involving topicality based on Corpus 3. The representation of this discourse function in English translation will also be discussed in this section.

7.4.1. Definiteness and referentiality

The use of numeral classifiers is closely related to establishing reference and definiteness/specificity, as they are less obligatory than gender in terms of their presence, choice, and ordering in classifier phrases (see in §2.4.2 and §3.3.2). Numeral classifiers tend to be compared to articles or determiners in non-classifier languages based on the definiteness or specificity they express. Furthermore, they can also indicate referentiality or identifiability of their referents. In this section, I will examine the two functions based on the different constructions they occur in.

Numeral classifiers expressing indefiniteness tend to occur in constructions without demonstratives. There are 481 noun phrases without demonstratives, compared with 164 noun phrases with demonstratives in Corpus 3. Among the 481 numeral phrases without demonstratives, there are 389 indefinite phrases mainly in two types of constructions: [NUM/PRON+CLF+(CLF)+N] and [CLF+N]. Example (109) shows indefinite phrases of the first type, where the numeral classifiers *míng* 'CLF:IDENTITY' and *fú* 'CLF:PICTURE' do not refer to any specific referents, only that *fú* 'CLF:PICTURE' is repeated to express a plural meaning.

174

(109) Indefiniteness expressed by numeral classifiers in Corpus 3

Zuòwéi	*yī*	*míng*	*fēngjǐng*	*shèyǐng*
as	one	CLF:IDENTITY	landscape	photography
àihàozhě,	*xiànshí*	*de*	*chǎngjǐng*	*jīngcháng*
enthusiast	real	MOD	scene	often
zài	*tā*	*yǎnzhōng*	*xíngchéng*	*yī*
in	his	eyes	form	one
fú	*fú*	*yìshù*	*gòutú*。	
CLF:PIC-TURE	CLF:PICTURE	art	composition	

'As **a landscape photography enthusiast**, Wang often saw the sights around him as **artistic compositions**.'

Numeral classifiers occurring in the construction of [CLF+N] also tend to express indefiniteness. According to Chen (2003: 1178-1179), numeral classifiers can be compared with definite determiners when they are introduced by the object marker *bǎ* or followed by proper nouns or kinship nouns (see §3.3.2). However, no examples of such use of numeral classifiers can be found in Corpus 3. Among the 23 occurrences of the construction of [CLF+N] in Corpus 3, the general classifier *gè* occurs in 22 examples and the only specific classifier used in this construction is *jiàn* 'CLF:PIECE'. The constructions of the type [CLF+N] tend to be introduced by verbs rather than the object marker *bǎ* and all of them are indefinite. As shown in example (110), the general classifier *gè* occurs in the structure [CLF+N] and indicates indefinite and non-specific reference.

(110) Indefiniteness expressed by [CLF+N] in Corpus 3

Tā	*yěxǔ*	*bú*	*shì*	*gè*
he	maybe	not	be	CLF:GENERAL
hǎo	*jǐngchá,*	*dàn*	*quèshí*	*shì*
good	cop	but	certainly	be
gè	*hěn*	*jiǎosè*。		
CLF:GENERAL	fearsome	role		

'Maybe he (Shi Qiáng) wasn't **a good cop**, but he was certainly **a fearsome one**.'

The indefiniteness expressed by numeral classifiers tends to be reflected in indefinite articles and determiners in English translation, as shown in the above two examples and in §6.2.2. When repeated numeral classifiers occur with the numeral *yī* 'one', the indefiniteness is shown in plural nouns, as in *artistic compositions* in example (109). However, the indefiniteness expressed by numeral classifiers can also be shown in bare nouns or nouns in plural forms in English translation even when the numeral *yī* 'one' is used with only

one numeral classifier. As shown in example (111), the indefinite reading of *yī* 'one' and the numeral classifier *gǔ* 'CLF:STRAND' is shown in the form of the bare noun, as *warmth* is an abstract and mass noun that cannot be directly used with an indefinite article.

(111) The form of bare nouns as representation of indefiniteness in English translation in Corpus 3

Wāng Miǎo	*de*	*xīnzhōng*	*yǒngqǐ*	*yī*
Wang Mioo	MOD	in.the.heart	fill	one
gǔ	*nuǎnliú*		...	
CLF:STRAND	warm.current		...	

'Wang (accepted the box), **warmth** filling his chest.'

The indefiniteness expressed by numeral classifiers can also be expressed in the plural form of nouns in English translation, even if there is only one numeral classifier present in the construction. As shown in example (112), *shòupí* 'animal.hide' in (a) is translated into *animal hides* due to the indefinite interpretation *yī* 'one' used with *zhāng* 'CLF:SPREADING OPEN/FLAT'. As discussed in §6.2.2, *yī* 'one' used with numeral classifiers can be equivalent to *each* or *every* in English so that the phrase appears in singular form but has a plural reading. This can be illustrated by the translation of the numeral noun phrase in (b) used with the numeral *yī* 'one' and the numeral classifier *bǎ* 'CLF:HANDLE' in the following clause. The numeral noun phrase is translated into a noun phrase in singular form, i.e., *a short, wide bronze sword*, to agree with the pronoun *each* in number in the sentence.

(112) The plural form of nouns as representation of indefiniteness in English translation in Corpus 3

a. ...	*tā*	*kàndào*	*liǎng*	*rén*
...	he	see	two	person
dōu	*shì*	*nánxìng,*	*pī*	*zhe*
both	be	male	drape	PROG
pòlàn	*de*	*chángpáo,*	*wàimiàn*	*hái*
ragged	MOD	robe	outside	also
guǒ	*zhe*	*yī*	*zhāng*	
wrap	PROG	one	**CLF:SPEADING OPEN/FLAT**	
āngzāng	*de*	*shòupí,*	...	
dirty	MOD	animal.hide	...	

'He saw that both figures were male. They were dressed in long robes full of holes, covered by **dirty animal hides**.'

b. ...   *dōu*   *dài*   *zhe*   ***yī***
   ...   both   carry   PROG   **one**

bǎ   ***qīngtóng***   ***shídài***   ***nà***
CLF:HANDLE   bronze   time   that

zhǒng   ***yòu***   ***kuān***   ***yòu***
CLF:KIND, GENERAL   also   wide   also

duǎn   *de*   ***jiàn***。
short   MOD   sword

'Each carried **a short, wide bronze sword**.'

Numeral classifiers can also be used to express definiteness. Numeral classifiers in 92 (out of 481) noun phrases without demonstratives are used to express definiteness. These noun phrases tend to be preceded by modifiers or be used anaphorically. As shown in example (113), the numeral noun phrase is preceded by modifiers *miànqián de* (in.front MOD) 'in front' and thus refers to the two specific officers.

(113) Definiteness expressed by numeral classifiers with modifiers in Corpus 3
   Wāng Miǎo   *bújiě*   *dì*   *kàn*   *zhe*   ***miànqián***
   Wang Miao   baffle   AUX   look   PROG   in.front

   de   ***liǎng***   ***wèi***   ***jūnguān***。
   MOD   two   CLF:INDIVIDUAL, RESPECT   army.officer

   'Wang looked at **the two officers**, baffled.'

Noun phrases used anaphorically also express definiteness. As shown in example (114), the numeral classifier *zuò* 'CLF:SEAT, PEDESTAL, BASE' in (a) refers anaphorically to the pendulums, while the two numeral noun phrases in (b) refer to the five experiments and the two balls that the speaker has mentioned previously in the novel. While the three phrases occur without demonstratives or determiners, they are also definite.

(114) Definiteness expressed by numeral classifiers for anaphora in Corpus 3

a. ...   *zhè*   *shēngyīn*   *shì*   *Cháogē*   *dàdì*
   ...   this   sound   be   Chaoge   land
   shàng   *xǔduō*   *qíguài*   *de*   *dōngxī*   *fāchū*
   over   many   strange   MOD   thing   generate
   de,   *nà*   *shì*   *yī*   *zuò*
   MOD   that   be   one   CLF:SEAT, PEDESTAL, BASE
   zuò   *jùdà*   *de*   *dānbǎi,*   **_měi_**
   CLF:SEAT, PEDESTAL, BASE giant   MOD   pendulum   each
   zuò   *dōu*   *yǒu*   *jǐshí*   *mǐ*
   CLF:SEAT, PEDESTAL, BASE all   have   tens   meter
   gāo。
   high

   'The sound was generated by the numerous giant pendulums that could be seen
   all over Zhao Ge, **each** tens of meters in height.'

b. ...   *zài*   **_wǔ_**   **_cì_**   **_shìyàn_**   *zhōng,*
   ...   at   five   CLF:TIME   experiment   in
   liǎng   **_gè_**   **_qiú_**   *de*   *zhìliàng*   *shì*
   two   CLF:GENERAL   ball   MOD   mass   be
   méiyǒu   *biànhuà*   *de*   ...
   no   change   MOD   ...

   'During **these five experiments**, the mass of **the two balls** never changed.'

All the 164 noun phrases with demonstratives are definite. Example (115) shows a definite use of numeral classifiers with demonstratives. In (a), the numeral classifier *wèi* 'CLF:INDIVIDUAL, RESPECT' is used with the demonstrative *nà* 'that' to specifically refer to the younger cop, while in (b), the general classifier *gè* is used with both the demonstrative *nà* 'that' and the numeral *liǎng* 'two'. The demonstrative *nà* expresses definite reference, while the numeral *liǎng* 'two' indicates a specific quantity.

(115) Definiteness expressed by numeral classifiers with demonstratives in Corpus 3

a. ...   *tā*   *jiù*   *xiàng*   *pángbiān*
   ...   he   then   toward   near
   _nà_   **_wèi_**   **_niánqīngrén_** *shìyì*
   that   CLF:INDIVIDUAL, RESPECT youngster   sign
   le   *yīxià,*   *hòuzhě*   *xiàng*   *Wáng Miǎo*
   PRT   one.time   latter   toward   Wang Miao
   chūshì   *le*   *jǐngguānzhèng*   ...
   show   PRT   badge   ...

   '...the man nodded at **the younger cop**, who showed Wang his badge...'

b. ...	*tā*	*jìng*	*zài*	*xiàyìshí*
...	he	unexpectedly	at	unconsciousness
zhōng	*shuōchū*	*le*	***nà***	***liǎng***
in	say	PRT	that	two
gè	***zhèshí***	***yīng***	***shífēn***	***jìhuì***
CLF:GENERAL	then	ought	completely	taboo
de	***míngcí***	...		
MOD	noun	...		

'Unconsciously, he had named **the two hypotheses that he ought to have avoided.**'

While the numeral *yī* 'one' is likely to be omitted when the noun phrase contains a demonstrative, it should be present when the quantity is highlighted in the context. As shown in example (116), the numeral classifier *gè* is used with both the demonstrative *nà* 'that' and the numeral *yī* 'one', with the demonstrative expressing definiteness and the numeral highlighting a particular seat.

(116) The presence of *yī* 'one' with demonstratives in Corpus 3

Tā	*búxiǎng*	*āi*	*Shǐ Qiáng*	*zuò,*
he	not.want	be.next	Shi Qiang	sit
dàn	*yě*	*zhīyǒu*	***nà***	***yī***
but	also	only	that	one
gè	*kōngwèi,*	*tā*	*zhīhǎo*	*zuò*
CLF:GENERAL	empty.seat	he	can.only	**sit**
guòqù	...			
over	...			

'He didn't want to sit next to Shi, but he had no choice, as that was **the only empty seat.**'

The results show that in noun phrases with both numeral classifiers and demonstratives, demonstratives rather than numeral classifiers are more likely to express definiteness, while numeral classifiers are more likely to specify referents based on their semantic features. Therefore, when numeral classifiers used with demonstratives are translated into English, the definite article *the* is a more direct equivalent of demonstratives than of numeral classifiers, while nouns are more likely to be translated into more specific ones in English, as shown by *the young cop* compared with *niánqīngrén* 'youngster' in (a), and *hypotheses* compared with *míngcí* 'nouns' in (b) in example (115).

Another construction more likely to be definite is based on the sequence of [NUM/DEM+N] with the omission of numeral classifiers. This structure is more likely

to refer to previously mentioned referents. As shown in the example (117), the noun *rén* 'person' in (a) is used with the numeral *yī* 'one' and refers to the proper noun *Wang Miao* that precedes immediately the construction and in (b) the noun *rén* 'person' is used with the demonstrative *zhè* 'this' and refers to *Ding Yi* mentioned in the first part of the sentence in (b).

(117) Definiteness expressed by numeral noun phrases without numeral classifiers in Corpus 3

a. ...
xiǎnrán	*tā*	*de*	*cūsú*
apparently	he	MOD	rough

búzhǐ	*shì*	*ràng*	*Wāng Miǎo*	*yī*
not.only	be	let	Wang Miao	one

rén	*bú*	*shìyīng*。
person	not	comfortable

'Apparently, Wang wasn't **the only one** annoyed by his rough manners.'

b.
Tā	*xiǎngdào*	*le*	*Dīng Yí,*	*kě*
he	think	PRT	Ding Yi	but

xiànzài	*zhè*	*rén*	*zìjǐ*	*yě*
now	this	person	self	also

xiànrù	*jīngshén*	*wēijī*	*zhīzhōng*。
sink	spirit	crisis	in

'He thought of Ding Yi, but **that man** was now in a spiritual crisis of his own.'

The construction [NUM/DEM+N] is similar to pronouns and expresses definiteness by replacing proper names when they are mentioned in subsequent contexts. In the English translation, such constructions are usually translated into phrases composed of the definite article and pronouns, as *the one* in (a), or demonstratives and nouns, as *that man* in (b).

The above examples have shown the use of numeral classifiers to indicate definiteness. A related way of the definite reading of numeral classifiers is to indicate referentiality or identifiability of their referents. While numeral classifiers in definite phrases refer to specific and thus identifiable referents, those found in indefinite phrases can indicate referentiality to different degrees, i.e., referentiality related to identifiable referents and nonreferentiality related to nonidentifiable referents or identifiable referents irrelevant to the discourse. Numeral classifiers used for indefinite but identifiable referents tend to occur in presentative or foregrounded structures, as illustrated in example (118).

(118) Identifiable nonspecific reference expressed by numeral classifiers in Corpus 3

a. *Huìyì*   *shì*   ***zài***   ***yī***   ***gè***
    meeting   be   at   one   CLF:GENERAL

    dàtīng   *lǐ*   *jǔxíng*   *de,*   *Wāng Miǎo*
    hall   in   hold   MOD   Wang Miao

    yī   *jìnqù*   *jiù*   *duì*   *zhèlǐ*
    once   enter   just   to   here

    de   *fēnluàn*   *chījīng*   *búxiǎo* 。
    MOD   chaos   surprise   rather

    'Wang was surprised by the chaos as he entered **the large meeting room**.'

b. ***yī***   ***míng***   ***shàoxiào***   ***jūnguān***   *jímáng*
    one   CLF:IDENTITY   major   army.officer   hurry

    shàngqián   ...
    step.forward   ...

    '**One of the army officers, a major**, stepped forward.'

As shown in example (118), in (a) the general classifier *gè* is used to introduce *dàtīng* 'hall' in the presentative structure *zài yī gè dàtīng* (at one CLF:GENERAL hall) 'in a hall', while in (b) *míng* 'CLF:IDENTITY' is used to introduce *shàoxiào jūnguān* 'major and army officer' in the foregrounded phrase *yī míng shàoxiào jūnguān* (one CLF:IDENTITY major army.officer) 'one army officer and major' as the subject in the clause. Their referents are mentioned for the first time and thus should be regarded as indefinite. However, both of the two referents are identifiable in the discourse and thus the two numeral classifiers are used for identifiable referentiality. In English translation, the identifiable reference expressed by numeral classifiers can be shown by the definite article, as *the* in *the large meeting room* in (a), or the pronoun, as *one* in *one of the army officers* in (b).

      Numeral classifiers used for nonreferentiality or low referentiality tend to occur as part of predicates. Example (119) shows nonreferential use of numeral classifiers as part of predicates.

(119) Numeral classifiers indicating nonreferentiality as part of predicates in Corpus 3

a. ...   *nà*   *shénqíng*   *jiù*   ***xiàng***
    ...   that   expression   just   be.like

    tā   *kàndào*   ***yī***   ***duǒ***   ***piāoliàng***
    she   see   one   CLF:FLOWER.LIKE   beautiful

    de   *yěhuā*   *yīyàng* 。
    MOD   wildflower   same

    '...the expression on her face was the same as when she saw **a pretty wildflower**.'

b. ...	*Kēxuébiānjiè*	*shì*	*yī*	*gè*
…	Science.frontier	be	one	CLF:GENERAL
yóu	*guójì*	*dǐngjiān*	*xuézhě*	*gòuchéng*
have	international	elite	scholar	constitute
de	*zǔzhī,*	*duì*	*tā*	*de*
MOD	organization	to	it	MOD
diàochá	*shì*	*yī*	*jiàn*	*jíqí*
investigate	be	one	CLF:PIECE	extremely
fùzá	*hé*	*mǐn'gǎn*	*de*	*shì。*
complex	and	sensitive	MOD	matter

'…the Frontiers of Science is (an organization) **made up of elite international scholars**. Investigating it is **an extremely complex and sensitive matter**.'

In example (119), there are three numeral noun phrases preceded by the copulas *xiàng* 'be.like' in (a) and *shì* 'be' in (b), i.e., *yī duǒ piāoliàng de yěhuā* (one CLF:FLOWER.LIKE beautiful MOD wildflower) 'a pretty wildflower' in (a) and *yī gè yóu guójì dǐngjiān xuézhě gòuchéng de zǔzhī* (one CLF:GENERAL have international elite scholar constitute MOD organization) 'an organization made up of elite international scholars' and *yī jiàn jíqí fùzá hé mǐn'gǎn de shì* (one CLF:PIECE extremely complex and sensitive MOD matter) 'an extremely complex and sensitive matter' in (b). The phrases do not refer to any specific referent and so the numeral classifiers express nonreferentiality. In English translation, nonreferential use is reflected in the indefinite article as shown in *a pretty wildflower* in (a), *an extremely complex and sensitive matter* in (b). Such nonreferential use may also be omitted in English translation as the referents concerned are irrelevant in the discourse, as shown in the translation of the first phrase in (b) with no mention of *an organization*.

Nonreferential use of numeral classifiers can also occur in the construction [CLF+N]. As shown in example (120), the general classifier occurs in the construction introduced by the copula *shì* 'be' and expresses nonreferentiality.

(120) Numeral classifiers indicating nonreferentiality in construction [CLF+N] in Corpus 3

tā	*zhēn*	*shì*	*gè*	*cōngmíng*	*rén。*
he	really	be	CLF:GENERAL	smart	man

'He's really **a smart man**.'

In summary, numeral classifiers can be used to manage reference related to definiteness and referentiality. They tend to express indefiniteness (389 out of 645) involving referentiality to different degrees rather than definiteness related to specific and identifiable

referents. Numeral classifiers used with numerals or without numerals in the construction of [CLF+N] tend to express indefiniteness, while those used with determiners or noun phrases used without numeral classifiers are more likely to be definite. The results show that Chinese numeral classifiers do not directly correspond to the article *the* in English, as numeral classifiers in pre-verbal phrases in the construction of [CLF+N] in Cantonese do (see §2.4.2). Instead, they tend to coerce an indefinite reading of numerals and express definiteness when used with determiners, typically demonstratives. In English translation, the function involving indefiniteness tends to be reflected in determiners or articles, with determiners and indefinite articles expressing indefiniteness and demonstratives or the definite article expressing definiteness. Numerals in the structures [DEM+NUM+CLF+(N)] and [NUM/DEM+N] are more likely to be directly translated into either their equivalent numerals or pronouns in English and nouns in these structures tend to be translated into more specific ones, as usually both the quantity and the reference that the phrases express are more highlighted.

As for referentiality, numeral classifiers in definite phrases also express identifiable and specific reference. In contrast, numeral classifiers in indefinite phrases can express referentiality when they are used in presentative structures and nonreferentiality when they are used as part of predicates. In English translation, referential use of numeral classifiers can be shown in the articles and pronouns, with the definite article or pronouns used to represent identifiable and specific reference while indefinite articles are used to express nonreferentiality. Unidentifiable referents may also be omitted in English translation, as they tend to be irrelevant to the discourse. The referential use of numeral classifiers is also related to the discourse function involving topicality, which will be discussed in §7.4.2.

7.4.2. Topicality

The use of Chinese numeral classifiers can also be analysed in terms of topicality related to the establishment and management of the status of referents in discourse. They can also be used to exhibit thematic salience of referents in different patterns. As discussed in §3.3.2, specific classifiers tend to be used to introduce new referents in foregrounded or presentative clauses, while the general classifier *gè* is likely to replace specific classifiers

to re-introduce referents in subsequent mentions in discourse. Numeral classifiers may highlight referents but not necessarily imply their thematic significance. This section will examine such uses of numeral classifiers and examine their representation in English in Corpus 3.

The use of numeral classifiers to introduce new referents and manage reference in subsequent mentions can be illustrated in the example (121). As shown in the example, two cops and two men in military uniform are introduced in the numeral noun phrases with the specific classifier *míng* 'CLF:IDENTITY', which is later replaced by the general classifier *gè* when the two men in military uniform are mentioned again in the second sentence. In English translation, the two pairs of referents are introduced by nouns. In terms of reference management, the two men in military uniform are referred to anaphorically by the pronominal phrase *the latter two* referring to the second of the two pairs of referents.

(121) Reference management in the construction of [NUM+CLF+N] in Corpus 3

Wāng Miǎo	*juédé,*	*lái*	*zhǎo*	*tā*
Wang Miao	think	come	find	him
de	*zhè*	*sì*	*gè*	*rén*
MOD	this	four	CLF:GENERAL	person
shì	*yī*	*gè*	*qíguài*	*de*
be	one	CLF:GENERAL	strange	MOD
zǔhé:	**liǎng**	**míng**	**jǐngchá**	*hé*
combination	two	CLF:IDENTITY	policeman	and
liǎng	**míng**	**jūnrén,**	*rúguǒ*	**nà**
two	CLF:IDENTITY	armyman	if	that
liǎng	**gè**	**jūnrén**	*shì*	*wǔjǐng*
two	CLF:GENERAL	armyman	be	armed.police
hái	*suàn*	*zhèngcháng,*	*dàn*	*zhè*
still	consider	normal	but	this
shì	*liǎng*	*míng*	*lùjūn*	*jūnguān。*
be	two	CLF:IDENTITY	army	officer

'Wang Miao thought the four people who came to find him made a rather odd combination: **two cops** and **two men in military uniforms**. If **the latter two** were armed police, that would be somewhat understandable, but they were actually PLA officers.'

Reference can also be managed in the construction of [NUM/DEM+N] without numeral classifiers. For example, Shiqiang, one of the main characters in the novel *Three-Body Problem*, is referred to eight times in noun phrases in such constructions as

[NUM+CLF+N] and [NUM+N] in Chapter 1. Four of the eight noun phrases are illustrated in examples (122) and (123).

(122) Reference management in the construction of [NUM+N] in Corpus 3

...	*dàn*	***nà***	***wèi***
…	but	that	CLF:INDIVIDUAL, RESPECT
biànyī	*jiù*	*ràng*	*rén*	*tǎoyàn*
plainclothes	just	let	people	detest
le.	***Zhè***	***rén***	*zhǎng*	*de*
PRT	this	person	grow	AUX
wǔdàsāncū,	*yīliǎnhèngròu,*		...
thickset	a.face.full.of.bulging.muscles	…

'…But **the other one, in plainclothes**, immediately grated on him. **He** was thickset and had a face full of bulging muscles.'

As shown in example (122), Shi Qiang is first referred to as one of the two policemen by the specific classifier *wèi* 'CLF:INDIVIDUAL, RESPECT' with the noun *biànyī* 'plainclothes', and then by the phrase *zhè rén* 'this person' without a numeral classifier. As discussed in §7.4.1, phrases composed of numerals or demonstratives and nouns have a similar function with pronouns. However, they tend to occur adjacent to nouns or noun phrases with the same reference. In English translation, such way of reference management tends to be represented by pronouns, as shown by the pronoun *he* in the example.

The general classifier *gè* tends to be used to reintroduce referents as a topic in later mentions in discourse. This is illustrated in example (123).

(123) Reintroduction of the referent as a topic in Corpus 3

Xiànzài,	*Wāng Miǎo*	*zhīdào*	*Cháng Wěisī*	*bǎ*
Now	Wang Miao	know	Chang Weisi	BA
tā	*yǐqián*	*de*	***zhè***	***gè***
he	previous	MOD	this	CLF:GENERAL
zhànshì	*diàolái*	*shì*	*yǒu*	*dàolǐ*
soldier	bring	be	have	reason
de,	***zhè***	***gè***	***wàibiǎo***	***cūsú***
MOD	this	CLF:GENERAL	appearance	vulgar
de	***jiāhuo,***	*yǎnjīng*	*gēn*	*dāozi*
MOD	guy	eye	as	knife
yīyàng。				
same				

'Now Wang Miao understood why it made sense for General Chang to have asked to have **this man who was once a soldier under his command**. **Shi**, who appeared so vulgar and careless, had eyes as sharp as knives.'

When a referent is reintroduced as a topic in discourse, numeral classifiers tend to be present in numeral noun phrases to avoid ambiguity and to allow the use of more specific nouns as well as more complex pre-nominal modifiers. As shown in example (123), Shi Qiang is mentioned again in the two numeral noun phrases with the general classifier *gè* and the nouns *zhànshì* 'soldier' and *jiāhuo* 'guy', with the latter preceded with the modifiers *wàibiǎo cūsú de* (appearance vulgar MOD) 'appearing vulgar'. In contrast, non-classified noun phrases tend to take fewer prenominal modifiers, according to Li (2000: 1120). In English translation, referents tend to be reintroduced by specific nouns or names postmodified with adjectives or clauses. As shown in the example (123), the two phrases are translated into a noun phrase *this man* and a proper name *Shi* followed by a non-defining relative clause.

General and specific classifiers can both occur in noun phrases as part of predicates. In such contexts, they do not refer to specific and identifiable referents; instead, they are used to ascribe a quality or characteristic to the referent, as discussed in §7.4.1. Examples of such use of numeral classifiers are shown in example (124).

(124) Numeral classifiers used to ascribe features to referents in Corpus 3

a. *Nǐ*   *zhème*   *shuōhuà*   *shízài*   *bú*
you   this.way   speak   really   not
xiàng   ***yī***   ***míng***   ***hégé***   ***de***
be.like   one   CLF:IDENTITY   qualified   MOD
jǐngguān 。
police.officer
'The way you speak is not appropriate for **a good police officer**.'

b. *Tā*   *yěxǔ*   *bú*   *shì*   ***gè***
he   maybe   not   be   CLF:GENERAL
hǎo   ***jǐngchá***,   *dàn*   *quèshí*   *shì*
good   policeman   but   certainly   be
gè   ***hěn***   ***jiǎosè*** 。
CLF:GENERAL   fearsome   role
'Maybe he wasn't **a good cop**, but he was certainly **a fearsome one**.'

As shown in the example (124), Shi Qiang is referred to in (a) by the numeral noun phrase based on the sequence of [NUM+CLF+N], where the specific classifier *míng* 'CLF:IDENTITY' is used by the speaker to remind Shi Qiang of his profession. In (b), Shi Qiang is described in the two phrases composed of the general classifier and nouns. In translation, the features or characteristics expressed in numeral noun phrases in Chinese are shown

in indefinite phrases in English, while the difference between specific and general classifiers is not reflected in the translation.

Numeral classifiers tend to be used to mark the thematic salience of referents, which is not necessarily related to thematic significance. They can be used to highlight referents in discourse by foregrounding noun phrases "for the purpose of vivifying or intensifying the description" Li (2000: 1118). However, the use of numeral classifiers does not indicate the importance of the referents as a topic in discourse. The use of numeral classifiers to mark thematic salience can be illustrated in example (125).

(125) The thematic salience marked by numeral classifiers in Corpus 3

Wáng Miǎo	*dìyī*	*yǎn*	*jiù*	*duì*
Wang Miao	first	eye	just	to
lái	*zhǎo*	*tā*	*de*	*jǐngchá*
come	find	him	MOD	policeman
méi	*yǒu*	*hǎogǎn.*	*Qíshí*	*nà*
not	have	good.impression	actually	that
míng	***chuān***	***jǐngfú***	***de***	***niánqīngrén***
CLF:IDENTITY	wear	police.uniform	MOD	youngster
hái	*xíng,*	*jǔzhǐ*	*hěn*	*yǒu*
still	fine	behavior	very	have
lǐmào,	*dàn*	***nà***	***wèi***	***biànyī***
politeness	but	that	CLF:INDIVIDUAL, RESPECT	plainclothes
jiù	*ràng*	*rén*	*tǎoyàn*	*le.*
just	let	people	detest	PRT

'As soon as Wang saw the cops, he felt annoyed. **The younger one** was all right—at least he was polite. But **the other one, in plainclothes**, immediately grated on him.'

While the younger cop is introduced by the numeral classifier *míng* 'CLF:IDENTITY' and the prenominal modifiers *chuān jǐngfú de* (wear police.uniform MOD) 'in police uniform', he is not thematically important and is not mentioned again in the discourse. In contrast, Shi Qiang, as one of the main characters in the novel, is of thematic significance in the discourse, although he is also introduced by the numeral noun phrase based on the construction [DEM+CLF+N] but without any prenominal modifiers. The thematic salience marked by numeral classifiers can be reflected in nouns or pronouns in English translation. As shown in the example (125), the two referents are referred to by subject phrases in the two sentences, with Shi Qiang referred to by the complex construction *the other one, in plainclothes*.

The results provide evidence for the claim made by Erbaugh (1986: 408) that a new referent is likely to be introduced by a specific classifier and subsequently referred to by a general classifier. However, specific classifiers still occur as part of predicates to ascribe qualities or characteristics to the referents and numeral classifiers can be omitted in noun phrases when the referents occur as subjects or immediately follow in the subsequent mentions. Furthermore, the presence of numeral classifiers can be used to foreground referents and allow more complex modifiers to precede nouns in numeral noun phrases. This discourse function of numeral classifiers tends to be reflected in the function of pronouns and nouns in English translation, where salient referents are likely to be referred to by subject phrases or nouns with post-modifiers.

7.5. Re-presentation of referents

The discourse function involving the re-presentation of referents involves a change of numeral classifiers for the same referent, as discussed in §2.4.2 and §3.3.2. There are only two examples related to this function in Corpus 3. This section will examine the use of numeral classifiers for the re-presentation of referents based on the two examples and also examine how this discourse function is reflected in English translation.

Numeral classifiers can be used to present referents from different perspectives. In Corpus 3, three different numeral classifiers are used with reference to Shi Qiang. Other than the function related to reference management, as discussed in §7.4.1, they also express different messages from different points of view. This is illustrated in example (126).

(126) The re-presentation of human referents with *zhǒng* in Corpus 3

    a. *…*   *dàn*   *nà*   *wèi*   *biànyī*

        …   but   that   CLF:INDIVIDUAL, RESPECT plainclothes

        jiù   *ràng*   *rén*   *tǎoyàn*   *le.*

        just   let   people   detest   PRT

        '…**the other one, in plainclothes**, immediately grated on him.'

    b. ***Zhè***   ***zhǒng***   ***rén***   *zěnme*

        this   CLF:KIND, GENERAL   person   how

        néng   *jìn*   *Zuòzhàn*   *zhōngxīn?*

        can   enter   battle   center

        'How can **a man like that** be part of the Battle Command Center?'

c.	...	**zhè**	**gè**	**wàibiǎo**	**cūsú**
	...	this	CLF:GENERAL	appearance	vulgar
	de	**jiāhuo,**	**yǎnjīng**	**gēn**	**dāozi**
	MOD	guy	eye	as	knife

yīyàng。
same

'**Shi**, who appeared so vulgar and careless, had eyes as sharp as knives.'

In (a), the numeral classifier *wèi* 'CLF:INDIVIDUAL, RESPECT' is used with the noun *biànyī* 'plainclothes' to introduce the referent Shi Qiang. In (b), Shi Qiang is referred to by the general kind classifier *zhǒng* with the noun *rén* 'person' by a major, and in (c), he is reintroduced by the general classifier *gè* and a more colloquial noun *jiāhuo* 'guy' from the perspective of Wang Miao again. While the specific classifier *wèi* in (a) expresses due respect to Shi Qiang as a policeman, the other two classifiers in (b) and (c) do not show enough respect for the referent. A certain degree of contempt is implied in (b), although *zhǒng* can used to refer to a type of people, as discussed in §5.3.4. In (c), the general classifier *gè* is more casual compared with *wèi* in (a) and thus shows that Wang Miao is less respectful toward Shi Qiang. In English, with the referents shown in nouns or pronouns, the affective meanings related to respect or contempt are not reflected in the translation.

The other example concerning re-presentation of referents is shown in example (127). In (a) Wang Miao is referred to by an officer as part of a group of scholars by *wèi* 'CLF:INDIVIDUAL, RESPECT' and *xuézhě* 'scholar'. However, in (b) he is referred to by Shi Qiang with the general classifier *gè*. As a professor, Wang Miao is respectable enough, and *wèi* 'CLF:INDIVIDUAL, RESPECT' is regarded as a more proper classifier for him typically in conversations. The general classifier, on the contrary, shows a certain degree of irony and contempt, especially when it is used with the modifier *zhùmíng* 'famous'.

(127) The representation of human referents in Corpus 3

a. *xiàwǔ*   *yǒu*   *yí*   *gè*   *zhòngyào*
afternoon   have   one   CLF:GENERAL   important
huìyì,   *yàoqǐng*   ***jǐ***   ***wèi***
meeting   invite   several   CLF:INDIVIDUAL, RESPECT
xuézhě   *hé*   *zhuānjiā*   *cānjiā,*   *shǒuzhǎng*
scholar   and   specialist   attend   general
ràng   *wǒmen*   *lái*   *yāoqǐng*   *nín*。
let   us   come   invite   you
'There's an important meeting this afternoon, to which **several scholars** and specialists are invited. The general sent us to invite you.'

b. ...   *xiàng*   *nǐ*   *zhèyàng*   ***yí***
…   like   you   such   one
gè   ***zhùmíng***   ***xuézhě****,*   *zǒnggāi*   *duì*
CLF:GENERAL   famous   scholar   should   to
gōnggòng   *ānquán*   *fùzé*   *ba*。
public   security   responsible   PRT
'…You're **a famous academic**. You have a responsibility toward the public welfare.'

In English translation, the noun *xuézhě* 'scholar' is translated into *scholars* and *academic*. While both *academic* and *scholar* are neutral in their connotations and can be interchangeable in most cases, the noun *academic* in (b) can imply that the referent is too theoretical from the perspective of Shi Qiang as an experienced policeman and thus the noun is relatively more derogatory compared with the noun *scholar*.

This section has examined the use of different numeral classifiers to recategorize the same referent. The results show that the change of numeral classifiers can present referents from different perspectives and express different affective messages concerning respect and contempt. With the omission of the numeral classifiers in English translation, the affective meanings can be shown in the choice of nouns, if they are not completely lost in translation.

7.6. Discussion and Concluding remarks

This section will summarize the findings related to the discourse functions of Chinese numeral classifiers and their representation in English translation based on Corpus 3. The results show that specific classifiers rather than general classifiers are more likely to be used for reference identification. Specific classifiers are more likely to be used for

anaphora and disambiguation and they tend to be used consistently in tracking referents to avoid ambiguity. In contrast, both general and specific classifiers can be used with demonstratives for deixis. Numeral classifiers used for reference identification tend to be represented by nouns or pronouns in English translation. Numeral classifiers for anaphora tend to be translated into such pronouns as *one* and *the other*, while those for deixis and disambiguation tend to be equivalent to more specific nouns in English translation. These findings partly support the hypothesis that the discourse function related to reference identification can be reflected in the choice of more specific nouns when they are used for deixis in English translation. However, when numeral classifiers are used for anaphora and disambiguation, they are not reflected in more specific nouns in English translation.

Regarding reference management, numeral classifiers can be used to express definiteness and referentiality and thus they tend to be compared with articles and determiners in non-classifier languages. However, the results show that indefiniteness tends to be expressed by numeral classifiers occurring in such classification structures as [NUM+CLF+(N)] and [CLF+N], while definiteness is more likely to be shown in numeral phrases, typically with demonstratives. The indefiniteness expressed by numeral classifiers can lead to different interpretations of the numeral *yī* 'one' as indefinite articles or determiners and as the form of bare nouns in English translation. In contrast, numerals used with demonstratives in classification structures or occurring in noun phrases without the presence of numeral classifiers tend to be more specific in quantity and be translated into their direct equivalents as numerals or pronouns in English.

Concerning the function related to referentiality, the referents of numeral classifiers in definite phrases tend to be identifiable and specific. In contrast, numeral classifiers in indefinite phrases can express referentiality to different degrees. They can express identifiable but nonspecific referents in presentative structures and nonreferentiality when they are used as part of predicates. The findings related to definiteness and referentiality support the hypothesis that the discourse function related to reference management can be shown in the use of articles in English. However, the results also show that numeral classifiers are more likely to be used with demonstratives to express definiteness in Chinese and *yī* 'one' can be interpreted differently as *one*, articles or plural due to the presence of numeral classifiers.

Reference management related to topicality can be shown in the introduction and re-introduction of referents by numeral classifiers in discourse and the thematic saliency

of referents. While numeral classifiers in general can be used to mark thematic salience in foregrounded or presentative classes, specific classifiers are more likely to be used to introduce new referents and the general classifier *gè* can be used to replace the specific classifiers to re-introduce the referents as a topic in subsequent discourse. In English translation, the discourse function related to topicality is represented by the use of nouns and pronouns, while salient referents tend to occur as subjects or nouns with post-modifiers.

The only two examples related to the discourse function of re-presentation of referents show that numeral classifiers can be used to present referents from different perspectives and express such affective meanings as respect or contempt. In English translation, this discourse function is either omitted or represented in the choice of nouns.

In summary, while the discourse function related to re-presentation of referents may be omitted in English translation, most discourse functions of Chinese numeral classifiers can be reflected in English translation in one way or another. Their discourse function of reference management involving definiteness and referentiality is reflected in articles and other determiners, including demonstratives, while the functions of reference identification related to anaphora, disambiguation and deixis, reference management involving topicality and even re-presentation of referents can be shown in the choice or the function of nouns or pronouns in English translation.

Chapter 8: Conclusions

8.1. Introduction

This chapter will conclude the study by summarizing the major research findings. As shown in the previous chapters, numeral classifiers have received substantial attention in linguistics. They differ from other types of nominal classification in terms of morphosyntactic and semantic features and related functions. They form with numerals a left-branching structure to quantify nouns. Furthermore, they denote semantic features of their referents. These lexical and grammatical features, as well as the related semantic and discourse functions, can be represented in corresponding elements when Chinese numeral classifiers are translated into English as a non-classifier language.

In this study, I have first examined the semantic contribution of Chinese numeral classifiers and related elements based on their frequency and their collocations with different types of nouns in the context with and without adjectives, based on two Chinese-English parallel corpora. I have also investigated the effect of the presence of numeral classifiers and the comparison of Chinese numeral classifiers with their English equivalents based on their functionality. The semantic functions of Chinese numeral classifiers and their representation in English have been examined in the two corpora based on numeral noun phrases with and without adjectives, while the discourse functions of Chinese numeral classifiers and their representation in English have been investigated in Corpus 3 derived from the novel *The Three-body Problem*.

The results have demonstrated the use of different types of numeral classifiers in Chinese and their graded degrees on the individuation hierarchy. Other than denoting semantic features of noun referents, numeral classifiers can also contribute additional

meanings related to quality and quantity to noun phrases. Furthermore, the study has also examined the functions of Chinese numeral classifiers based on the three corpora and shown that Chinese numeral classifiers can be compared with such elements in English as measure words, the plural form, articles and determiners, nouns and pronouns to fulfil their different semantic and discourse functions.

In this chapter, the findings related to the semantic contribution of numeral classifiers will first be discussed in §8.2. Then Chinese numeral classifiers will be compared with English equivalent elements based on their functionality in §8.3. The study will be concluded by suggesting issues for further research on Chinese numeral classifiers in §8.4.

8.2. Semantic contribution of Chinese numeral classifiers

The data from Corpus 1 and Corpus 2 have shown that different types of numeral classifiers contribute to the semantics of the noun phrases in different ways. This section will discuss the complementary use of different types of Chinese numeral classifiers and their semantic correlation with such elements as nouns and adjectives in §8.2.1. The semantic contribution of numeral classifiers will be analysed in §8.2.2, while §8.2.3 will concern more general issues regarding the nature of classification of classifiers and numeral noun phrases.

8.2.1. The use of Chinese numeral classifiers and the individuation hierarchy

In Chapter 2 and Chapter 3, I reviewed different ideas people have about the uses of general and specific classifiers and the semantics of Chinese numeral classifiers based on their three types: entity, event and kind. In Chapter 5, based on the collected data in Corpus 1 without adjectives, I showed that different types of numeral classifiers are used in a complementary way in the context without adjectives. The general classifiers, typically *gè* 'CLF:GENERAL' (46.43%), are used far more frequently than the remaining more than 100 specific classifiers (40.07%), which complement general classifiers by expressing more salient features of noun referents.

Based on the collocations of numeral classifiers with different types of nouns, numeral classifiers can be interpreted in terms of individuation hierarchy based on their referents. As shown in Table 37, numeral classifiers for humans and animals are ordered as the top two as they are the most identifiable ones, numeral classifiers for bounded entities, events and concepts are ranked in the middle, and numeral classifiers for mass nouns used with reference to entities that are fuzzy in their physical boundedness are ranked lowest on the hierarchy.

Table 37. Individuation hierarchy and semantic correlation of Chinese numeral classifiers

individuation hierarchy		general classifiers	specific classifiers
Human		*gè*	entity classifiers for humans
Animal		*zhī*	entity classifiers for animals and shape classifiers
Bounded	entity	*gè*	entity classifiers
inani-	event	*gè*	event classifiers
mate	concept	*gè*	less specified entity classifiers and shape classifiers
Neutralized entity		*gè, zhǒng*	entity, event, and kind classifiers
Mass		*zhǒng*	kind classifiers

The three general classifiers tend to be used as default classifiers to show different degrees of individuation: *gè* 'CLF:GENERAL' for human and nonhuman bounded entities, *zhī* 'CLF:SINGLE' for animals, and *zhǒng* 'CLF:KIND, GENERAL' for uncountable referents. Specific entity and event classifiers tend to be ordered on the top or in the middle on the hierarchy. For example, specific human classifiers are ranked on the top for human referents, followed by specific animal classifiers and other specific classifiers for inanimates.

There is also variation in the use of numeral classifiers with nouns ranked higher or lower on the individuation hierarchy. For example, the general classifier *gè* as a default classifier for human and inanimate nouns can be used with nouns for animals, and then, animal referents used with *gè* are upgraded on the hierarchy and attributed with human qualities. In contrast, the general kind classifier *zhǒng* as a default classifier for uncountable nouns can also be used with countable nouns to refer to a kind or a type of noun referents.

In Corpus 2 with adjectives, the results showed that there are two-way restrictions of numeral classifiers and adjectives based on their semantic preference. The general classifiers rarely occur with adjectives, while specific classifiers, typically shape classifiers, are used predominantly in this context. On the other hand, numeral classifiers also restrict the choice of adjectives. Adjectives preceding numeral classifiers tend to denote size and

shape. Therefore, numeral classifiers and adjectives in this context have a semantic preference for shape and size and denote semantics related to both quality and quantity. The quantity reading is typical of measure words. Therefore, numeral classifiers used with adjectives are more like measures words and can be used with a wider range of nouns compared with those used in the context without adjectives.

These results confirm that general classifiers *gè* is used far more frequently than other numeral classifiers in Chinese and thus it is usually regarded as the only general classifier in Chinese (e.g., Erbaugh 1986: 402; Aikhenvald 2000: 324; Gao and Malt 2009: 1177), while the other two general classifiers *zhī* 'CLF:SINGLE' and *zhǒng* 'CLF:KIND, GENERAL' occur more frequently with nouns for animals and uncountable nouns, respectively. However, what should be clarified is that the general classifier *gè* is rarely used with nouns for animals and uncountable nouns or in the context with adjectives. The results thus corroborate the claims about the limitations in the use of *gè* with nouns for animals (Zhou 2014: 91; Frankowsky and Ke 2016: 63-65). In the context with adjectives, the use of general and specific classifiers shows a different scenario. Specific classifiers, typically shape classifiers, are far more frequently used with adjectives, while all the three general classifiers *gè*, *zhī* 'CLF:SINGLE' and *zhǒng* 'CLF:KIND, GENERAL' rarely occur with adjectives.

The results also provide evidence for the semantic correlation between numeral classifiers and other elements in numeral noun phrases, based on the distribution of general and specific classifiers on the individuation hierarchy and semantic preference shown in the collocation of numeral classifiers with adjectives. The results thus show that the choice of numeral classifiers depends on the semantics of their head nouns as well as the presence of adjectives, and all the elements in numeral noun phrases can contribute to the semantics of the phrases concerned. This issue will be further discussed in the following sections.

8.2.2. Semantic contribution of numeral classifiers to noun phrases

This section will deal with the semantic contribution of numeral classifiers. I will first discuss the semantic correlation of numeral classifiers and nouns based on the

individuation hierarchy and then turn to the new meanings that numeral classifiers express when they are used with nouns ranked on different levels on the individuation hierarchy.

Numeral classifiers denote semantic features of noun referents either in a broad sense or in a more specific way. Like other types of nominal classification, general classifiers in Chinese also have a semantic core related to humanness and animacy, although their semantics are not transparent and they are used with broad categories of nouns. In contrast, specific classifiers complement general classifiers by denoting more specific features of noun referents, as shown above in §0 in the distribution of general and specific classifiers on the individuation hierarchy. For example, specific human classifiers, e.g., *wèi* 'CLF:INDIVIDUAL, RESPECT' and *míng* 'CLF:IDENTITY', are ranked on the top on the individuation hierarchy for human referents, while specific animal classifiers, e.g., *pí* 'CLF:HORSE, HORSE-LIKE ANIMALS' and *tóu* 'CLF:HEAD', tend to specify salient features of animal referents, and specific entity classifiers, e.g., *kuài* 'CLF:LUMP.LIKE' and *liàng* 'CLF:VEHICLE, CAR', tend to occur with bounded inanimates to denote their physical features.

Numeral classifiers can also contribute new meanings to noun phrases. They tend to attribute new properties to noun referents when they are used with nouns ordered on a different level on the individuation hierarchy. For instance, when the general classifier *gè* is used with nouns for animals, it attributes human properties to animal referents, as shown in the above section. Numeral classifiers can also contribute new meanings to noun phrases when they are used with adjectives. As shown in the above section, shape classifiers are most frequently used in the context of adjectives of size and shape. Shape classifiers and adjectives can express both physical features of shape and quantity related to size of referents. Therefore, shape classifiers with adjectives can express new meanings related to quantity and quality to referents and are more like measure words by being used with a wider range of nouns without necessarily denoting the features of noun referents.

The results confirm that most numeral classifiers can be distinguished from measure words based on their semantic properties, as discussed in §2.4.1.1 and §3.2.3. Numeral classifiers and nouns are semantically correlated. However, when the use of numeral classifiers involves the reassignment of nouns on the individuation hierarchy and the presence of adjectives, numeral classifiers may contribute extra meanings related to both quality and quantity to noun referents, as mentioned above. This finding disproves the semantic criterium concerning the applicability of numeral classifiers to more limited

groups of nouns (Ahrens 1994: 204), as shape classifiers used with adjectives are more like measure words and can be applied to a wider range of nouns, as shown above. Furthermore, it should be noted that *gè*-substitution proposed by Wang (1994: 27-36) cannot test all numeral classifiers, typically when it involves the use of numeral classifiers with uncountable nouns and nouns for animals. The classifier *gè* hardly occurs with uncountable nouns, while substituting animal classifiers with *gè* may change the interpretation of referents with the attributed human properties.

In conclusion, numeral classifiers can contribute semantics to noun phrases by being semantically correlated with their head nouns or by contributing additional meanings related to quality and quantity to noun phrases. The results confirm that numeral classifiers denote semantic features of noun referents. Furthermore, the results also show that numeral classifiers can contribute new meanings to noun phrases when they are used with nouns ranked on a different level on the individuation hierarchy or when they are used with adjectives. The two types of semantic contribution of numeral classifiers are related to the semantics of numeral noun phrases, which will be discussed in the following section.

8.2.3. The interpretation of noun phrase reference

Numeral classifiers classify referents instead of nouns, as all Chinese numeral classifiers, general or specific, have a semantic core related to humanness and animacy. However, it is difficult to determine to what degree numeral classifiers contribute to noun phrase reference and to interpret the semantics of the phrases as a whole (cf. Lucy 2000: 335-337; Senft 2012: 10-13). The interpretation of the reference of noun phrases can be examined based on the collocation of numeral classifiers and nouns and the related semantic contribution of numeral classifiers as well as other elements in numeral noun phrases.

Different types of numeral classifiers are used with their corresponding types of nouns based on their shared semantics on the individuation hierarchy, as shown above in §0. For example, human classifiers are used with human nouns, animal classifiers are used with nouns for animals, other entity classifiers are used with nouns for inanimates, event classifiers are used with nouns for events, and kind classifiers cooccur with mass

nouns. Based on this type of the collocation, the referents can be interpreted in a broad sense as humans, animals or inanimate entities.

The more fixed noun-classifier pairing involves the use of some specific classifiers with specific nouns or a group of specific nouns. Some numeral classifiers are collocated with a small number of nouns based on their overlapping semantics. They include such specific classifiers as *běn* 'CLF:BOOK', *pǐ* 'CLF:HORSE, HORSE-LIKE ANIMAL', *piān* 'CLF:ARTICLE', *liàng* 'CLF:VEHICLE, CAR', and *sōu* 'CLF:SHIP'. While referents can be interpreted directly from the numeral classifiers, nouns express more obvious and specific meanings, as shown in the use of *běn* 'CLF:BOOK' with *shū* 'book', *jiàocái* 'textbook' and *zìdiǎn* 'dictionary'. Therefore, numeral classifiers in this type are referential while nouns may contribute more transparent or even additional information to noun phrase reference.

Other numeral classifiers occurring in fixed collocations with nouns are specific classifiers denoting salient features. They tend to cooccur with a certain type of nouns based on their semantics. These numeral classifiers include *jiā* 'CLF:HOUSEHOLD' for nouns with reference to organizations and institutions, *kē* 'CLF:PLANT' and *zhū* 'CLF:STALK' for plants, and *dòng* 'CLF:BUILDING, BEAM' and *zhuàng* 'CLF:BUILDING, PILLAR' for buildings (see §5.3.3.3), as well as such event classifiers as *cháng* 'CLF:VENUE' and *cì* 'CLF:TIME' for nouns denoting events. The semantics of these numeral classifiers are also transparent. However, both numeral classifiers and nouns contribute to the meaning of noun phrases, as shown in the collocation of *kē* 'CLF:PLANT' with *shù* 'tree' and *cǎo* 'grass'. These specific classifiers tend to occur with countable concrete nouns (see §5.3.3.3) and nouns for events (see §5.3.3.4). They are less likely to be replaced by general classifiers, typically when they are used with countable concrete nouns, according to Zhou (2014: 91). Therefore, this type of numeral classifiers contributes to noun phrase reference by highlighting salient features of noun referents.

Some numeral classifiers are less fixed in the collocation without nouns, as they are applied to more varied groups of nouns. These numeral classifiers include the general classifiers, the less specified entity classifiers, e.g., *jiàn* 'CLF:PIECE', and shape classifiers. The semantics of the former two types of numeral classifiers tend to be opaque, other than the semantic core related to humanness and animacy, and thus they tend to make distinctions only among such referents as humans, animals, or inanimates. As regards shape classifiers, they tend to be used with wider range of nouns with reference to concrete entities and denote transparent features related to shape, size and dimension. For example,

one of the shape classifiers *tiáo* 'CLF:SLENDER' can be assigned to various types of nouns with reference to animals, e.g., dogs, concrete entities, e.g,, roads and scarves, as well as abstract concepts, e.g., news. The classifier conveys a long and slender shape, either physically or conceptually, and are thus referential. Numeral classifiers in this type also contribute to noun phrase reference by specifying salient features, while the semantics of related noun phrases should be interpreted based on all the elements, including numeral classifiers, nouns and even adjectives. As shown in example (128), the shape classifier *piàn* 'CLF:FLAT/THIN' is used with the noun *lántiān* 'blue.sky' in (a) to refer to the sky, while in (b), the shape classifier *tiáo* 'CLF:SLENDER' is used with the adjective *cháng* 'long' and the noun *tiānkōng* 'sky' to refer to a long stretch of the sky.

(128) The interpretation of noun phrase reference
  a. *yī*   *piàn*   *lántiān*
    one   CLF:FLAT/THIN   blue.sky
    'the sky'
  b. *yī*   *cháng*   *tiáo*   *tiānkōng*
    one   long   CLF:SLENDER   sky
    'a long stretch of the sky'

To conclude, the more fixed collocation of numeral classifiers with nouns, the greater degree to which they contribute to noun phrase reference. Those specific classifiers in fixed pairing with a limited number of nouns are the most referential by directly referring to referents, while nouns collocated with these types of nouns may provide additional information to the referents of noun phrases. Most specific classifiers denoting salient features can be collocated with a certain type of nouns and be referential by highlighting the specific features of their referents. Similarly, shape classifiers can refer to referents by denoting their shape, size and dimension, although they can be collocated with more varied types of nouns. The least referential numeral classifiers are general classifiers and less specified entity classifiers. These numeral classifiers are the most flexible in terms of their collocation with different types of nouns. However, they can also make distinctions in a broad sense among humans, animals and inanimate entities. The results also show that numeral classifiers tend to co-contribute with nouns and adjectives to the reference of noun phrases. These findings confirm that numeral classifiers in Chinese are referential, as discussed in §3.2.2, and can contribute to "adequate noun phrase reference" (Lucy 2000: 329), as discussed in §2.2. However, what should be noted is that the

referents of noun phrases should be interpreted based on all elements in the phrases, rather than solely numeral classifiers, nouns or adjectives.

8.3. Chinese numeral classifiers vs. English equivalent elements

In §3.2.1 and §3.2.2, I have analysed the syntactic and semantic features of Chinese numeral classifiers. In §2.4.2 and §3.3, I have also discussed the functions of classifiers in general and the functions of Chinese numeral classifiers in particular. Based on these analyses, I further examined Chinese numeral classifiers based on three corpora and discussed how their lexical and grammatical features as well as semantic and discourse functions are reflected in English translation. This section will discuss the effect of the presence and absence of numeral classifiers by comparing them with their equivalent elements in English translation based on their semantic and discourse functions. In §8.3.1, Chinese numeral classifiers will be compared with such elements as measure words, the plural form, and nouns in English based on their semantic functions. In §8.3.2, numeral classifiers will be compared with such elements as the articles and determiners, and nouns and pronouns in English based on their discourse functions.

8.3.1. Comparison based on semantic functions

The data from Corpus 1 and Corpus 2 have shown that the semantic functions of Chinese numeral classifiers are rarely reflected in English translation, since numeral classifiers tend to be omitted. The direct equivalents of Chinese numeral classifiers are measure words, as they occur in similar structures between numerals and nouns in the two languages. Only 3.06% of numeral classifiers are equivalent to measure words in English translation in Corpus 1 without adjectives, while 58.13% of numeral classifiers are equivalent to measure words in Corpus 2 with adjectives. What should also be noted is that numeral classifiers in general are far less likely to be pre-modified with adjectives. Furthermore, among different types of numeral classifiers, the general classifier *gè* is rarely (20 out of 3111) translated into measure words in English in Corpus 1. The results support the hypothesis that Chinese numeral classifiers tend to be omitted in English translation.

In the following sections, I will examine how equivalent English elements are used to represent the semantic functions of Chinese numeral classifiers if they are not totally omitted in the translation.

8.3.1.1. Chinese numeral classifiers vs. English measure words

Measure words in English translation reflect two major semantic functions of numeral classifiers related to individuation of nouns and differentiating referents. First, English measure words correspond to Chinese numeral classifiers to create a semantic unit to quantify uncountable nouns. Data showed that numeral classifiers are significantly more likely to be equivalent to measure words when they are used with uncountable nouns in Corpus 1 (20.75% with uncountable nouns vs. 1.94% with countable nouns) and Corpus 2 (80% with uncountable nouns vs. 52.86% with countable nouns). The result provides support for the hypothesis that semantic units created by numeral classifiers tend to be reflected in measure words in English when their head nouns are uncountable.

Second, English measure words may also reflect the function of differentiating referents of Chinese numeral classifiers. While most entity classifiers are equivalent to such general measure words as *piece*, and thus, their specific properties are not reflected in English translation, some entity classifiers can be equivalent to English measure words and express specific properties, as shown in the comparison of *céng* 'CLF:LAYER' with *layer* and *coat* in English translation. It should be noted that the choice of measure words in English can be equivalent to nouns or modifiers of nouns rather than numeral classifiers in Chinese. As shown in example (129), the measure word *stretch* is translated based on the noun *shānqū* 'mountain.area' instead of the general classifier *gè* in (a), and the measure word *fragment* is derived from the head noun *suìpiàn* 'fragment' instead of the numeral classifier *kuài* 'CLF:LUMP.LIKE' in (b).

(129) The choice of measure words based on nouns instead of numeral classifiers
    a. *yī*   *gè*   *shānqū*
       one   CLF:GENERAL   moutain.area
       'a stretch of mountains'
    b. *yī*   *xiǎo*   *kuài*   *fúdòng huósāi suìpiàn*
       one   small, little CLF:LUMP.LIKE floating piston *fragment*
       'a fragment of floating piston'

Data in the two corpora also showed that measure words in English translation can reflect the meanings that numeral classifiers contribute to noun phrases, other than the lexical meanings related to specific properties of noun referents as discussed above. These measure words tend to be equivalent to numeral classifiers used with nouns ranked either higher or lower on the individuation hierarchy, as entity classifiers used with uncountable nouns or kind classifiers used with countable nouns. Some entity classifiers can also be equivalent to English measure words in the translation when numeral classifiers are used with countable nouns ranked on the same level on the individuation hierarchy. In this case, numeral classifiers tend to express both specific features of their referents and quantity related to size, as shown in the use of (*xiǎo* 'small') *kuài* 'CLF:LUMP.LIKE' with the noun *dàngāo* 'cake' to refer to *a (small) piece of cake*. The results showed that a small proportion of measure words in English can represent the function of differentiating referents of numeral classifiers as well as the lexical meanings that numeral classifiers contribute to noun phrases.

8.3.1.2. Chinese numeral classifiers vs. English plural marker

Other than measure words, the other form in English used to represent the function of individuation of nouns is the plural marker. The results demonstrated that most nouns in English translation in Corpus 1 and Corpus 2 are shown in singular form with indefinite articles (over 82.78% in Corpus 1 and 80.69% in Corpus 2) and other determiners (12.51% in Corpus 1 and 17.58% in Corpus 2), as they are equivalent to Chinese nouns in phrases based on the sequence of [*yī* 'one'+CLF+N] and [*yī* 'one'+ADJ+CLF+N]. However, it is worth mentioning that the indefinite reading of *yī* 'one' can sometimes be reflected in the plural form in English nouns, as in this case, *yī* 'one' is equivalent to the English determiner *every* or *each*, which appears in a singular form but has a plural reference. While uncountable nouns need to be individuated by measure words in English, countable nouns can be directly enumerated. However, the plural suffix is added to nouns when they have a plural reading in English. The results corroborate the argument that numeral classifiers and plural makers are in the same grammatical category, only that plural markers are omittable when nouns are shown in singular form (e.g., T'sou 1976; Borer 2005; Cowper and Hall 2012; Doetjes 2011; Her 2012; Mathieu 2012), as discussed in §2.4.2 and §3.3.

The results related to the translation of *yī* 'one' also support the hypothesis that the numeral *yī* 'one' is more likely to be translated into articles expressing definiteness rather than the numeral *one* expressing specificity.

8.3.1.3. Chinese numeral classifiers vs. English nouns

Only a few nouns in English in the corpus correspond with numeral classifiers to show functions related to differentiating referents and ascribing properties to referents. Most nouns in English translation do not show different properties expressed by different numeral classifiers, as shown in *writer* as an equivalence of *zuòjiā* 'writer' whether used with the general classifier *gè* or the specific classifiers *wèi* 'CLF:INDIVIDUAL, RESPECT' and *míng* 'CLF:IDENTITY', and *elephant* as an English equivalent of *dàxiàng* 'elephant' whether used with the general animal classifier *zhī* 'CLF:SINGLE' and the specific classifier *tóu* 'CLF:HEAD'. Only a few nouns in English translation reflect the specific properties expressed by numeral classifiers when their head nouns are polysemous or unspecific in Chinese, as shown in the different choice of nouns *beard*, *moustache* and *soul patch* in opposition to the numeral classifiers *fù* 'CLF:SET, ATTITUDE', *bǎ* 'CLF:HANDLE', and *kuài* 'CLF:LUMP.LIKE' with *húzi* 'beard, moustache'. Therefore, the results contradict the hypothesis that specific properties denoted by Chinese numeral classifiers tend to be expressed by nouns in English.

As regards the function related to ascribing properties to referents, it is rarely shown in nouns in English translation either. The properties attributed to noun referents tend to be omitted in English translation, as shown in the translation of nouns for animals with human classifiers and animal classifiers and human nouns collocated with the general classifier *gè* and the specific classifier *wèi* 'CLF:INDIVIDUAL, RESPECT'. Only a few examples in Corpus 1 showed that nouns in English translation reflect the affective meanings denoted by numeral classifiers, as shown in *homebody* instead of *housewife* as a translation of *jiātíng zhǔfù* 'family housewife' used with the general kind classifier *zhǒng* 'CLF:KIND, GENERAL'. The results, therefore, provide support to the hypothesis that semantic properties attributed to noun referents by Chinese numeral classifiers tend to be omitted in English translation.

In conclusion, the semantic functions of Chinese numeral classifiers are rarely reflected in English translation. In those limited examples with English equivalent elements used to represent the functions of numeral classifiers, English measure words are equivalent to numeral classifiers to show the functions of individuating uncountable nouns and differentiating referents, while the plural marker -*s* in English is equivalent to numeral classifiers to individuate countable nouns. A few nouns in English represent the functions of differentiating referents and ascribing properties to referents, although most nouns in English are direct equivalents of nouns in Chinese and do not show the specificity expressed by numeral classifiers.

8.3.2. Comparison based on discourse functions

The discourse functions of classifiers in general were discussed in §2.4.2 and the discourse functions of Chinese numeral classifiers in particular were examined in §3.3.2. In Chapter 7, I analysed the use of Chinese numeral classifiers in discourse based on Corpus 3, including such questions as how and to what degree Chinese numeral classifiers are used to identify and manage reference and recategorize referents in discourse and how these functions are represented in English translation. In this section, I will compare Chinese numeral classifiers with English articles and determiners in §8.3.2.1 and nouns and pronouns in §8.3.2.2.

8.3.2.1. Chinese numeral classifiers vs. English articles and determiners

Chinese numeral classifiers can be compared with articles and determiners in English to express definiteness and referentiality in relation to the function of reference management. Data in Corpus 3 showed that Chinese numeral classifiers are more likely to express indefiniteness (389 out of 645). They can either coerce an indefinite reading of numerals that they occur with or express indefiniteness directly in the construction of [CLF+N] typically involving the use of the general classifier *gè* introduced by verbs. Indefiniteness tends to be represented by indefinite articles, determiners or even bare nouns in English translation. For example, the numeral *yī* 'one' is more likely to be translated into the

indefinite articles under the influence of numeral classifiers, while bare nouns in English translation are usually shown as abstract or mass nouns. Indefiniteness can also be shown in the plural form of nouns in English based on the translation of the numeral *yī* 'one' with repeated numeral classifiers or even just one numeral classifier, as *yī* 'one' used with numeral classifiers can be equivalent to *every* or *each* in English which appears in a singular form but have a plural reading.

Data in Corpus 3 also showed a tendency of the use of numeral classifiers with determiners, typically demonstratives, to express definiteness. All the 164 noun phrases used with demonstratives are definite, while only 92 (out of 481) noun phrases without demonstratives are used to express definiteness. The 92 noun phrases without demonstratives tend to occur with modifiers or be used anaphorically, which will be discussed in §8.3.2.2. While the definite reading of these noun phrases tends to be shown in the definite article in English translation, the definite article is more likely to be equivalent to demonstratives rather than numeral classifiers in Chinese. In contrast, numeral classifiers and numerals in the construction of [DEM+(NUM)+CLF+N] are more likely to express specificity concerning quantity and quality of referents. For example, in the phrase *nà yī gè kòngwèi* (that one CLF:GENERAL empty.seat) 'the only empty seat', the definite article is equivalent to the demonstrative *nà* 'that' and *only (one)* is equivalent to the numeral *yī* 'one' to highlight and specify the quantity. Similarly, in the phrase *nà wèi niánqīngrén* (that CLF:INDIVIDUAL, RESPECT youngster) 'the young cop', the definite article is also equivalent to the demonstrative *nà* 'that' and *wèi* 'CLF:INDIVIDUAL' is used to specify the referent and thus the phrase is translated into *the young cop* instead of *the young man*. Phrases based on the sequence of [NUM/DEM+N] also have a definite reading, but they are more similar to pronouns in English and are usually translated into phrases composed of the definite article and nouns or pronouns, as shown in the phrases *the one* or *that man*, which will be further discussed in §8.3.2.2.

Articles in English translation can also be used to reflect the referentiality of numeral classifiers in discourse. The definite article can represent identifiable and specific reference of numeral classifiers in definite phrases, while the indefinite article can be used to express nonreferentiality to different degrees. Data in Corpus 3 showed that numeral classifiers used for indefinite but identifiable referents tend to occur in presentative or foregrounded structures, while those used for nonreferential or referents with low referentiality tend to occur as part of predicates. While identifiable referents can be represented

by such pronouns as *one* in English (see §8.3.2.2.), unidentifiable referents tend to be represented only by the indefinite articles or even omitted when they are irrelevant in the discourse.

The results confirm that Chinese numeral classifiers can correspond with determiners to express definiteness (e.g., Cheng and Sybesma 2012a; Li and Bisang 2012; Li and Wu 2018) and provide evidence for the hypothesis that definiteness expressed by Chinese numeral classifiers tend to be expressed by articles in English. However, the results also showed that in the context of demonstratives, numerals and numeral classifiers are more likely to express specificity in terms of quantity and quality of referents while demonstratives are more likely to directly express definiteness. Furthermore, the results showed that Chinese numeral classifiers do not occur in pre-verbal phrases based on the sequence of [CLF+N], as numeral classifiers in Cantonese do (Cheng and Sybesma 2012b). Therefore, Chinese numeral classifiers do not correspond directly to the article *the* in English when they express definiteness.

8.3.2.2. Chinese numeral classifiers vs. English nouns and pronouns

In English translation, nouns and pronouns are more likely to be used to represent the discourse functions of numeral classifiers concerning reference management related to topicality, reference identification, and re-presentation of referents. In this section, Chinese numeral classifiers will be compared with English nouns and pronouns in terms of the discourse functions other than reference management related to definiteness and referentiality.

Data in Corpus 3 showed that Chinese numeral classifiers can be used to establish and manage the status of referents in discourse in several ways. Specific classifiers tend to be used to introduce referents and are substituted by the general classifier *gè* when the referents are reintroduced in discourse. Numeral classifiers can even be omitted in the construction of [NUM/DEM+N] and the numeral noun phrases without numeral classifiers, similar to English pronouns, can be used to refer to previously mentioned referents. Furthermore, both general and specific classifiers can occur in noun phrases as part of predicates to ascribe qualities or characteristics to referents. In comparison, in English translation, referents tend to be introduced by nouns and referred to again by pronouns.

In later mentions, nouns tend to be used to re-introduce referents in English. However, the features or characteristics expressed in numeral noun phrases in Chinese tend to be shown in indefinite phrases in English, as discussed in §8.3.2.1, while the differences between specific and general classifiers in these structures are not reflected in the translation. Thematic salience of referents, which is not necessarily related to thematic significance, can be marked by numeral classifiers in presentative structures, which may take more prenominal modifiers. Such use of numeral classifiers tends be shown in English translation by nouns occurring as subjects or with post-modifiers.

Nouns and pronouns in English can also be compared with Chinese numeral classifiers in terms of reference identification related to anaphora, deixis and disambiguation. Data in Corpus 3 showed that numeral classifiers are more likely to be used for anaphora (52 out of 645) than for deixis (3 out of 645) or disambiguation (6 out of 645). Specific classifiers rather than general classifiers are more likely to occur without the presence of nouns for anaphora and disambiguation, while both general and specific classifiers can be used with demonstratives for deixis. In English translation, pronouns and nouns are more likely to be used for reference identification. Pronouns, e.g., *one* and *the other*, tend to be used for anaphora, while specific nouns tend to be equivalent to numeral classifiers for deixis and disambiguation. However, the results also showed that the consistent use of numeral classifiers to track referents and the specificity expressed by numeral classifiers for anaphora are not reflected in pronouns in English translation.

English nouns can be used to show the discourse function of numeral classifiers concerning re-presentation of referents. While there are only two examples of numeral classifiers related to this function, the results did show that numeral classifiers can be used to present referents from different perspectives and express affective messages. In English translation, with the omission of numeral classifiers, this function is either completely lost or reflected in the choice of nouns, as shown in the use of *academic*, implying a derogatory massage as being too theoretical, in opposition to *scholar*.

The results provide evidence for most hypotheses related to the discourse functions of Chinese numeral classifiers and their representation in English translation. First, specific classifiers are more likely to be used for reference identification related to anaphora and disambiguation. However, both general and specific classifiers can be used for deixis and the re-presentation of referents in discourse. Second, concerning reference management related to topicality, referents are likely to be introduced by specific

classifiers and referred to by the general classifier *gè* in later mentions, which also corroborates the argument by Erbaugh (1986: 408). However, it should be clarified that it is not the only way to track referents in discourse. For example, specific classifiers can still occur in discourse for the same referents, as they are not likely to be replaced by other classifiers, including the general classifier *gè*, when they are used for anaphora and disambiguation. Both specific and general classifiers can be used in noun phrases as part of predicates to ascribe properties to referents concerned. The construction [NUM/DEM+N] without the presence of numeral classifiers can function as pronouns to refer anaphorically to referents when the noun phrases occur as subjects or immediately follow nouns with the same reference. When numeral classifiers are used to mark thematic salience of referents, both general and specific classifiers can be used in presentative structures and allow more complex prenominal modifiers to occur in numeral noun phrases. Thirdly, the results provided evidence that numeral classifiers are likely to be represented by nouns in English to express discourse functions of reference identification and recategorization of referents. However, it is worth mentioning that English pronouns are more likely to represent the function of reference identification related to anaphora.

In conclusion, the discourse functions of Chinese numeral classifiers are more likely to be reflected in equivalent elements in English translation, compared with the degree of the representation of their semantic functions, as discussed in §8.3.1. Articles and determiners in English can be equivalent to numeral classifiers to express definiteness and referentiality, while the choice and function of English nouns and pronouns can reflect the discourse functions of numeral classifiers concerning reference management related to topicality, reference identification and re-presentation of referents. However, some discourse functions of numeral classifiers are still completely lost in English translation. For example, the specificity shown in numeral classifiers tends to be omitted when pronouns, e.g., *one* and *the other*, are used to show the function concerning anaphora. Similarly, the differences between specific and general classifiers are not shown in English translation when they are used in noun phrases as part of predicate to ascribe features or characteristics to referents. Numeral classifiers occurring in such structures may be omitted as they refer to referents irrelevant in discourse. Finally, the recategorization of referents may not be shown in English translation with the omission of numeral classifiers and nouns translated into their equivalents.

8.4. Prospects for further studies

While this study has reached some conclusions regarding the semantic contribution and English representation of the functions of Chinese numeral classifiers, it also presented some topics that need to be further investigated. First, more corpus-based studies are required to examine Chinese numeral classifiers in more varied constructions of noun phrases and types of genres in discourse. While this study has been based on three corpora, there are some limitations. For example, the source of data is not provided and numeral noun phrases are limited to [*yī* 'one' +(ADJ)+CLF+N] in Corpus 1 and Corpus 2. In Corpus 3, data is only derived from one type of genre, fiction, and the size of the corpus is significantly smaller compared with that of Corpus 1. The semantic and grammatical features of Chinese numeral classifiers should be further examined based on the interplay of numeral classifiers with other elements, e.g., demonstratives and numerals other than *yī* 'one' in numeral noun phases, while the extent and discourse functions of numeral classifiers typically related to re-presentation of referents should be examined on a larger corpus based on more types of genres.

Further research should also be conducted to compare numeral classifiers with nouns or pronouns in Chinese in terms of such functions as reference management, reference identification and the re-presentation of referents. These functions of numeral classifiers tend to be represented by nouns and pronouns in English translation, and numeral classifiers are not the only means to realize these functions in Chinese. Therefore, it is necessary to examine the extent and variability of numeral classifiers in opposition to those of nouns and pronouns based on their respective discourse functions. More comparative work should also be done on numeral classifiers in Chinese and types of nominal classification in other languages, including both classifier and non-classifier languages. Other topics that deserve attention in these comparative studies include the presence and absence of numeral classifiers in relation to the size of the nominal lexicon and the complexity of prenominal modifiers. Comparative work should also be done on numeral classifiers in translation. For example, numeral classifiers can be compared in Chinese as a source language and as a target language. They may show different degrees of variability in the choice of numeral classifiers and of the presentation of different functions.

The cognitive production and processing of reference identification and management in contexts with and without numeral classifiers might also prove an important area

for further studies. Numeral classifiers are semantically correlated with nouns and there are two-way restrictions on the choice of numeral classifiers. Furthermore, they can coerce an indefinite reading of numerals and are thus comparable with indefinite articles in non-classifier languages. Further studies should thus deal with how reference is identified and processed in discourse under the effect of the presence and absence of numeral classifiers in classifier and non-classifier languages.

Future investigations are also necessary to compare numeral classifiers in Mandarin Chinese with those in other varieties of Chinese languages. These languages vary in terms of the size and inventory of numeral classifiers, their semantic and syntactic features, and functions typically related to reference management. It can be interesting to investigate numeral classifiers in these languages and compare the results with those in Mandarin Chinese, or other classifier languages.

Abstract

This dissertation aims at a corpus-based analysis of Chinese numeral classifiers, elements occurring between numerals or demonstratives and nouns, in the translation between Mandarin Chinese and English. While Chinese numeral classifiers have received substantial attention in recent decades, relatively little research has been done on their semantic contribution based on corpus data and their representation in English translation, particularly regarding their functionality. In this study I show how Chinese numeral classifiers and other elements of the noun phrase contribute to its semantics and how they correspond to equivalent elements in English translation based on their semantic and discourse functions. Both quantitative and qualitative studies were conducted to address the above issues based on three self-compiled specialized Chinese-English parallel corpora with 6700 pairs of Chinese-English numeral noun phrases without adjectives, 523 pairs of Chinese-English numeral noun phrases with adjectives, and 645 pairs of noun phrases derived from five chapters of the novel *The Three-Body Problem* by Liu Cixin, respectively.

Based on the frequency and collocations of Chinese numeral classifiers, this study shows that different types of classifiers are used in a complementary way and can be ordered in terms of different degrees on the individuation hierarchy. I also show that numeral classifiers contribute additional meanings related to quality and quantity to noun phrases, other than denoting semantic features of noun referents. Moreover, I demonstrate that classifiers can be referential to different degrees based on how fixed they are in collocations with nouns, and the meaning of a classifier phrase is a product of its constituent parts.

With regard to the translation of Chinese numeral classifiers, the study shows that to express equivalent meanings, English uses such means as measure words, the plural

form, articles and demonstratives, nouns and pronouns. The study demonstrates that the semantic functions of Chinese numeral classifiers are rarely reflected in English translation. For example, the function involving individuation of nouns is only reflected in English in a limited number of measure words used with uncountable nouns or the plural form marked on countable nouns. The semantic functions concerning differentiating referents and ascribing properties to referents are only reflected in English translation in some measure words and nouns. The study also shows that compared with their semantic functions, discourse functions of Chinese numeral classifiers are more likely to be reflected in English translation. English articles and demonstratives can be used to express definiteness and referentiality related to the function of reference management, while nouns and pronouns can be also used to represent the functions related to reference management as well as reference identification and re-presentation of referents.

These findings contribute to a better theoretical and empirical understanding of Chinese numeral classifiers. This study demonstrates that general classifiers are used as default classifiers to categorize referents in terms of humanness and animacy, while specific classifiers are used to denote specific features of referents. In addition, this study illustrates that the general classifier *gè* is used to replace specific classifiers in subsequent mentions of a referent (reference management), while specific classifiers are used to track referents anaphorically and to disambiguate among previously mentioned referents. The results demonstrate the different degrees of grammaticalization of general and specific classifiers, and thus contribute to the typology of nominal classification in general. Furthermore, the study demonstrates the effect of the presence of numeral classifiers on the choice of nouns and adjectives and on the interpretation of the numeral *yī* 'one' as a marker of indefiniteness. While numeral classifiers are semantically correlated with nouns, there are two-way restrictions on the choice of numeral classifiers and adjectives. Numeral classifiers also coerce an indefinite reading of numerals. These results inform empirical studies on the cognitive mechanisms of language processing in relation to discourse functions of classifiers in Chinese, or more generally, classifier languages. Finally, I also demonstrate the different forms used in English translation to express meanings of Chinese numeral classifiers and their functions. Such contrastive analyses shed new light on the acquisition of Chinese numeral classifiers and equivalent forms in English by speakers of both non-classifier and classifier languages.

Streszczenie

Niniejsza rozprawa ma na celu badanie korpusowe chińskich klasyfikatorów liczbowych, elementów występujących między liczebnikami lub zaimkami wskazującymi a rzeczownikami, w tłumaczeniu między jęz. chińskim (mandaryńskim) a jęz. angielskim. Podczas gdy klasyfikatorom liczbowym w jęz. chińskim poświęcono w ostatnich dziesięcioleciach znaczną uwagę, przeprowadzono stosunkowo niewiele badań nad ich funkcjami semantycznymi w oparciu o dane korpusowe i ich reprezentacją w tłumaczeniu na język angielski, szczególnie w odniesieniu do ich funkcjonalności. W niniejszym opracowaniu pokazuję, w jaki sposób chińskie klasyfikatory liczbowe i inne elementy frazy nominalnej przyczyniają się do jej znaczenia i do jakiego stopnia odpowiadają one ekwiwalentnym elementom w tłumaczeniu na język angielski w oparciu o ich funkcje semantyczne i pragmatyczne. W celu rozwiązania powyższych problemów przeprowadzono zarówno badania ilościowe, jak i jakościowe, w oparciu o trzy samodzielnie utworzone wyspecjalizowane chińsko-angielskie korpusy równoległe zawierające 6700 par chińsko-angielskich fraz rzeczownikowych bez przymiotników, 523 par chińsko-angielskich fraz rzeczownikowych liczebnikowych z przymiotnikami oraz 645 par wyrażeń rzeczownikowych pochodzących z pięciu rozdziałów powieści *The Three-Body Problem* autorstwa Liu Cixin.

Opierając się na częstości występowania i kolokacjach chińskich klasyfikatorów liczbowych, niniejsze badanie pokazuje, że różne typy klasyfikatorów używane są w sposób komplementarny i mogą być uporządkowane pod względem różnych stopni w ramach 'hierarchii indywiduacji'. Pokazuję również, że klasyfikatory liczbowe wnoszą dodatkowe znaczenia ilościowe i jakościowe do fraz rzeczownikowych, inne niż oznaczanie cech semantycznych desygnatów nominalnych. Co więcej, pokazuję, że klasyfikatory mogą być referencyjne w różnym stopniu w zależności od tego, jak bardzo są utrwalone

w kolokacjach z rzeczownikami, oraz że na znaczenie wyrażenia klasyfikującego mają wpływ znaczenia jego części składowych.

Jeśli chodzi o tłumaczenie chińskich klasyfikatorów liczbowych, badanie pokazuje, że do wyrażenia równoważnych znaczeń w jęz. angielskim używane są takie elementy, jak kwantyfikatory, liczba mnoga, przedimki i wyrażenia wskazujące, rzeczowniki oraz zaimki. Badanie pokazuje, że funkcje semantyczne chińskich klasyfikatorów liczbowych rzadko znajdują odzwierciedlenie w tłumaczeniu na język angielski. Na przykład, funkcja polegająca na indywidualizacji rzeczowników jest odzwierciedlona w języku angielskim tylko w ograniczonej ilości kwantyfikatorów używanych z rzeczownikami niepoliczalnymi lub w liczbie mnogiej w przypadku rzeczowników policzalnych. Funkcje semantyczne dotyczące różnicowania desygnatów i przypisywania im cech afektywnych znajdują odzwierciedlenie w tłumaczeniu na język angielski tylko w niektórych kwantyfikatorach i rzeczownikach. Badanie pokazuje również, że w porównaniu z funkcjami semantycznymi, funkcje dyskursywne chińskich klasyfikatorów liczbowych mają większe prawdopodobieństwo na odzwierciedlenie w tłumaczeniu na język angielski. Angielskie przedimki i wyrażenia wskazujące mogą być używane do wyrażania określoności i referencyjności związanych z funkcją 'reference management', podczas gdy rzeczowniki i zaimki mogą być również używane do wyrażania funkcji związanych z 'reference management' a także 'reference identification' oraz 're-presentation of referents'.

Wyniki te przyczyniają się do lepszego teoretycznego i empirycznego zrozumienia chińskich klasyfikatorów liczbowych. Niniejsze badanie pokazuje, że klasyfikatory ogólne są używane jako klasyfikatory domyślne do kategoryzowania referentów jako osobowe vs. bezosobowe lub ożywione vs. nieożywione, podczas gdy klasyfikatory specyficzne służą do oznaczania cech specyficznych referentów. Ponadto badanie to pokazuje, że klasyfikator ogólny *gè* zastępuje klasyfikatory specyficzne w kolejnych odniesieniach do danego referenta ('reference management'), podczas gdy klasyfikatory specyficzne są używane do anaforycznego śledzenia referencji i do ujednoznaczniania pomiędzy wcześniej wymienionymi referentami. Wyniki pokazują różne stopnie gramatykalizacji klasyfikatorów ogólnych i specyficznych, a tym samym przyczyniają się do badań nad typologią klasyfikacji nominalnej. Ponadto w badaniu wykazano wpływ obecności klasyfikatorów liczebnikowych na dobór rzeczowników i przymiotników oraz na interpretację liczebnika *yī* 'jeden' jako wyznacznika nieokreśloności. Pomimo tego, że

klasyfikatory liczbowe są semantycznie skorelowane z rzeczownikami, istnieją dwukierunkowe ograniczenia w wyborze klasyfikatorów liczbowych i przymiotników. Klasyfikatory liczbowe wymuszają również nieokreśloną interpretację liczebników. Wyniki te przyczyniają się do badań empirycznych nad kognitywnymi mechanizmami przetwarzania języka w odniesieniu do funkcji pragmatycznych klasyfikatorów w języku chińskim oraz w innych językach posiadających klasyfikatory. Ilustruję również różne formy używane w tłumaczeniu na język angielski w celu wyrażenia znaczeń chińskich klasyfikatorów liczbowych i ich funkcji. Te analizy kontrastywne rzucają nowe światło na akwizycję chińskich klasyfikatorów liczbowych i ekwiwalentnych form w języku angielskim przez użytkowników zarówno języków nieposiadających i posiadających klasyfikatory.

References

Adams, Karen L., Alton L. Becker and Nancy Faires Conklin. 1975. Savoring the differences among classifier systems. (Paper presented at the Eighth International Conference on Sino-Tibetan Languages and Linguistics, University of California, Berkeley, 24–25, October 1975.).

Adams, Karen L. and Nancy Faires Conklin. 1973. "Toward a theory of natural classification", in: Claudia Corum, T. Cedric Smith-Stark and Ann Weiser (eds.), *Papers from the Ninth Regional Meeting, Chicago Linguistic Society, April 13–15, 1973*. Chicago: Chicago Linguistic Society, 1-10.

Ahrens, Kathleen. 1994. "Classifier production in normals and aphasics", *Journal of Chinese Linguistics* 22, 2: 202-247.

Aikhenvald, Aleksandra Y. 2000. *Classifiers: A typology of noun categorization devices*. Oxford: Oxford University Press.

Aikhenvald, Aleksandra Y. 2004. "Nominal classification: Towards a comprehensive typology", *STUF – Language Typology and Universals* 57, 2/3: 105-116.

Aikhenvald, Aleksandra Y. 2012. "Reflecting the world around us: Genders, noun classes, and classifiers", in: Alexandra Y. Aikhenvald (ed.), *The languages of the Amazon*. Oxford: Oxford University Press, 279-303. (https://doi.org/10.1093/acprof:oso/9780199593569.003.0010) (date of access: 08 Jan.2019).

Aikhenvald, Aleksandra Y. 2016a. *How gender shapes the world*. Oxford: Oxford University Press.

Aikhenvald, Aleksandra Y. 2016b. "The rise and fall of linguistic genders", in: Alexandra Y. Aikhenvald (ed.), *How gender shapes the world*. Oxford Scholarship Online.

(https://doi.org/10.1093/acprof:oso/9780198723752.003.0006) (date of access: 28 Mar. 2022).

Aikhenvald, Aleksandra Y. 2017. "A typology of noun categorization devices", in: Alexandra Y. Aikhenvald and R. M. W. Dixon (eds.), *The Cambridge handbook of linguistic typology.* (Cambridge Handbooks in Language and Linguistics.) Cambridge: Cambridge University Press, 361-404.

Alcock, Rutherford. 1861. *Elements of Japanese grammar, for the use of beginners.* London: Trübner.

Alexiadou, Artemis, Liliane Haegeman and Melita Stavrou. 2007. *Noun phrase in the generative perspective.* Berlin and New York: Mouton De Gruyter.

Allan, Keith. 1977. "Classifiers", *Language* 53, 2: 285-311.

Allassonnière-Tang, Marc and Marcin Kilarski. 2020. "Functions of gender and numeral classifiers in Nepali", *Poznań Studies in Contemporary Linguistics* 56, 1: 113-168.

Audring, Jenny. 2006. "Pronominal gender in spoken Dutch", *Journal of Germanic Linguistics* 18, 2: 85-116.

Audring, Jenny. 2008. "Gender assignment and gender agreement: Evidence from pronominal gender languages", *Morphology* 18: 93-116.

Audring, Jenny. 2017. "Calibrating complexity: How complex is a gender system?", *Language Sciences* 60: 53-68.

Barz, Richard K. and Anthony V. N. Diller. 1985. "Classifiers and standardisation: Some South and South-East Asian comparisons", in: David Bradley (ed.), *Language policy, language planning and sociolinguistics in South-East Asia.* Canberra: Pacific Linguistics, 155-184.

Becker, Alton L. 1975. "A linguistic image of nature: The Burmese numerative classifier system", *Linguistics* 165: 109-121.

Becker, Alton L. 1986. "The figure a classifier makes: Describing a particular Burmese classifier", in: Colette Craig (ed.), *Noun classes and categorization.* Amsterdam: John Benjamins, 327-343.

Beckwith, Christopher I. 2007. *Phoronyms: Classifiers, class nouns, and the pseudopartitive construction.* New York: Peter Lang.

Bi, Yanchao, Xi Yu, Jingyi Geng and F.-Xavier Alario. 2010. "The role of visual form in lexical access: Evidence from Chinese classifier production", *Cognition* 116, 1: 101-109.

Bisang, Walter. 1996. "Areal typology and grammaticalization: Processes of grammaticalization based on nouns and verbs in East and Mainland South East Asian languages", *Studies in Language* 20, 3: 519-597.

Bisang, Walter. 2014. "Overt and hidden complexity: Two types of complexity and their implications", *Poznań Studies in Contemporary Linguistics* 50, 2: 127-143.

Blank, David L. 1982. *Ancient philosophy and grammar: The syntax of Apollonius Dyscolus*. Chico, CA: Scholars Press.

Blank, David L. (ed.). 1996. *Ammonius: On Aristotle, on interpretation 1–8*. London: Duckworth.

Borer, Hagit. 2005. *Structuring sense*. Vol. 1: *In name only*. Oxford: Oxford University Press.

Brinton, Daniel G. 1885. "On polysynthesis and incorporation as characteristics of American languages", *Proceedings of the American Philosophical Society* 23: 48–86.

Brusciottus, [Hyacinthus] a Vetralla. 1659. *Regulae quaedam pro difficillimi congensium idiomatis faciliori captu ad grammaticae normam redactae*. Roma: Propaganda Fide.

Bu, Lian-zeng. 2011a. "南方汉语中量词定指现象来源初探——以桂南平话为例 [On the source of identifiability of classifiers: Take Guinan Pinghua as an Example]", *Studies in Language and Linguistics* 31, 3: 95-102.

Bu, Lian-zeng. 2011b. "汉语名量词起源再探 [A new study on the origin of noun classifiers]", *Journal of Jinan University (Philosophy and Social Sciences)* 150, 01: 89-96, 164.

Chao, Yuan Ren. 1968. *A grammar of spoken Chinese*. Berkeley, CA: University of California Press.

Chen, Ping 2003. "Indefinite determiner introducing definite referent: A special use of 'yi 'one'+classifier' in Chinese", *Lingua* 113: 1169-1184.

Cheng, Lisa L., Caroline Heycock and Roberto Zamparelli. 2017. "Two levels of definiteness", in: Michael Yoshitaka Erlewine (ed.), *Proceedings of GLOW in Asia XI, vol. 1*. (MIT Working Papers in Linguistics #84.) Cambridge, MA: MIT Working Papers in Linguistics.

Cheng, Lisa L. and Rint Sybesma. 1998. "*yi-wan tang, yi-ge tang*: Classifiers and massifiers", *Tsing-Hua Journal of Chinese Studies* 28, 3: 385-412.

Cheng, Lisa L. and Rint Sybesma. 1999. "Bare and not-so-bare nouns and the structure of NP", *Linguistic Inquiry* 30, 4: 509-542.

Cheng, Lisa L. and Rint Sybesma. 2012a. "Classifiers and DP", *Linguistic Inquiry* 43, 4: 634-650.

Cheng, Lisa L. and Rint Sybesma. 2012b. "Classifiers in four varieties of Chinese", in: Guglielmo Cinque and Richard S. Kayne (eds.), *The Oxford handbook of comparative syntax*. (https://doi.org/10.1093/oxfordhb/9780195136517.013.0007) (date of access: 31 Aug. 2015).

Chien, Yu-Chin, Barbara Lust and Chi-Pang Chiang. 2003. "Chinese children's comprehension of count-classifiers and mass-classifiers", *Journal of East Asian Linguistics* 12, 1: 91-120.

Chierchia, Gennaro. 1998a. "Plurality of mass nouns and the notion of 'semantic parameter'", in: Susan Rothstein (ed.), *Events and grammar*. Amsterdam: Kluwer Academic Publishers, 53-103.

Chierchia, Gennaro. 1998b. "Reference to kinds across language", *Natural Language Semantics* 6, 4: 339-405.

Coblin, W. South and Joseph A. Levi (eds.). 2000. *Francisco Varo's grammar of the Mandarin language (1703): An English translation of 'Arte de la lengua Mandarina'*. Amsterdam: John Benjamins.

Contini-Morava, Ellen. 2002. "(What) do noun class markers mean?", in: Wallis Reid, Ricardo Otheguy and Nancy Stern (eds.), *Signal, meaning, and message: Perspectives on sign-based linguistics*. Amsterdam: John Benjamins, 3-64.

Contini-Morava, Ellen and Eve Danziger. 2018. "Non-canonical gender in Mopan Maya", in: Sebastian Fedden, Jenny Audring and Greville G. Corbett (eds.), *Noncanonical gender systems*. Oxford: Oxford University Press, 129-146.

Contini-Morava, Ellen and Marcin Kilarski. 2013. "Functions of nominal classification", *Language Sciences* 40: 263-299.

Corbett, Greville G. 1991. *Gender*. Cambridge: Cambridge University Press.

Corbett, Greville G. 2000. *Number*. Cambridge: Cambridge University Press.

Corbett, Greville G. 2006. "Grammatical gender", in: Keith Brown (ed.), *Encyclopedia of language & linguistics*. Amsterdam: Elsevier, 749-756.

Corbett, Greville G. 2009. "Agreement", in: Sebastian Kempgen, Peter Kosta, Tilman Berger and Karl Gutschmidt (eds.), *Die slavischen Sprachen[The Slavic Languages]*. Berlin: De Gruyter, 342-354.

Corbett, Greville G. 2013a. "Number of genders", in: Matthew S. Dryer and Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (http://wals.info/chapter/30) (date of access: 14 Jul. 2019).

Corbett, Greville G. 2013b. "Systems of gender assignment", in: Matthew S. Dryer and Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (http://wals.info/chapter/32) (date of access: 22 May 2021).

Corbett, Greville G. and Sebastian Fedden. 2016. "Canonical gender", *Journal of Linguistics* 52, 3: 495–531.

Corbett, Greville G., Sebastian Fedden and Raphael Finkel. 2017. "Single versus concurrent systems: Nominal classification in Mian", *Linguistic Typology* 21, 2: 209-260.

Cowper, Elizabeth and Daniel Currie Hall. 2012. "Aspects of individuation", in: Diane Massam (ed.), *Count and mass across languages*. Oxford: Oxford University Press, 27-53.

Craig, Colette. 1986a. "Jacaltec noun classifiers: A study in language and culture", in: Colette Craig (ed.), *Noun classes and categorization*. Amsterdam: John Benjamins, 263-293.

Craig, Colette (ed.). 1986b. *Noun classes and categorization*. Amsterdam: John Benjamins.

Crosthwaite, Peter, Yuk Yeung, Xuefei Bai, Li Lu and Yeonsuk Bae. 2018. "Definite discourse-new reference in L1 and L2: The case of L2 Mandarin", *Studies in Second Language Acquisition* 40, 3: 625-649.

Crystal, David. 2008. *A dictionary of linguistics and phonetics*. Oxford: Blackwell.

Danziger, Eve and Ellen Contini-Morava. 2020. "Referential anchoring without a definite article: The case of Mopan (Mayan)", in: Balogh Kata, Latrouite Anja and D. Van Valin Robert, Jr. (eds.), *Nominal anchoring: Specificity, definiteness and article systems across languages*. Berlin: Language Science Press, 81-114.

Denny, J. Peter. 1976. "What are noun classifiers good for?", in: Salikoko S. Mufwene, Carol A. Walker and Sanford B. Steever (eds.), *Papers from the Twelfth Regional Meeting, Chicago Linguistic Society, April 23–25, 1976*. Vol. 12. Chicago: Chicago Linguistic Society, 122-132.

Ding, Shengshu. [1961] 1999. *现代汉语语法讲话* [Modern Chinese grammar]. Beijing: The Commercial Press. [Originally published, Beijing: The Commercial Press.].

Dixon, Robert M. W. 1972. *The Dyirbal language of North Queensland*. Cambridge: Cambridge University Press.

Dixon, Robert M. W. 1982a. "Classifiers in Yidiny", in: Robert M. W. Dixon (ed.), *Where have all the adjectives gone?* Berlin: De Gruyter Mouton, 185-205.

Dixon, Robert M. W. 1982b. "Noun classes", in: Robert M. W. Dixon (ed.), *Where have all the adjectives gone? And other essays in semantics and syntax*. Berlin: De Gruyter Mouton, 157-184.

Dixon, Robert M. W. 1982c. *Where have all the adjectives gone? And other essays in semantics and syntax*. Berlin: De Gruyter Mouton.

Dixon, Robert M. W. 1986. "Noun classifiers and noun classification in typological perspective", in: Colette Craig (ed.), *Noun classes and categorization*. Amsterdam, Philadelphia: John Benjamins, 105-112.

Dixon, Robert M. W. 1988. *A grammar of Boumaa Fijian*. Chicago: University of Chicago Press.

Doetjes, Jenny. 2011. "Count/mass distinctions across languages", in: Claudia Maienborn, Klaus von Heusinger and Paul Portner (eds.), *Semantics: An International Handbook of Natural Language Meaning*. Berlin: De Gruyter, 2259-2580.

Dong, Fanghua and Xiaohua Deng. 2019. "个体量词的跨语言新视野 [New cross-linguistic perspective on count-noun classifier]", *Studies in Language and Linguistics* 39, 1: 1-6.

Drapeau, Lynn and Renée Lambert-Brétière. 2011. "Verbal classifiers in Innu", *Anthropological Linguistics* 53, 4: 1-30.

Ebrey, Patricia Buckley. 1996. *The Cambridge illustrated history of China*. Cambridge: Cambridge University Press.

Emeneau, Murray Barnson. 1956. "India as a lingustic area", *Language* 32, 1: 3-16.

Enfield, Nicholas J. 2004. "Nominal classification in Lao: A sketch", *STUF – Language Typology and Universals* 57, 2: 117–143.

England, Nora C. 1983. *A grammar of Mam, a Mayan language*. Austin, TX: University of Texas Press.

Erbaugh, Mary S. 1986. "Taking stock: The development of Chinese noun classifiers historically and in young children", in: Colette Craig (ed.), *Noun classes and categorization*. Amsterdam: John Benjamins, 399-437.

Erbaugh, Mary S. 2006. "Chinese classifiers: The use and acquisition", in: Ping Li, Li Hai Tan, Elizabeth Bates and Ovid J.L. Tzeng (eds.), *Handbook of East Asian psycholinguistics: Chinese*. Cambridge: Cambridge University Press, 39-51.

Erbaugh, Mary S. 2013. "Classifier choices in discourse across the seven main Chinese dialects", in: Zhuo Jing-Schmidt (ed.), *Increased empiricism-recent advances in Chinese linguistics*. Amsterdam: John Benjamins, 101-125.

Fedden, Sebastian, Jenny Audring and Greville G. Corbett (eds.). 2018. *Non-canonical gender systems*. Oxford: Oxford University Press.

Fedden, Sebastian and Greville G. Corbett. 2017. "Gender and classifiers in concurrent systems: Refining the typology of nominal classification", *Glossa: A Journal of General Linguistics* 2, 1: 34,1-47.

Feng, Shengli. 1997. "Prosodic structure and compound words in Classical Chinese", in: Jerome Lee Packard (ed.), *New approaches to Chinese word formation: Morphology, phonology and the lexicon in modern and ancient Chinese*. 197-260.

Feng, Shengli. 2012. "The syntax and prosody of classifiers in Classical Chinese", in: Dan Xu (ed.), *Plurality and classifiers across languages in China*. Mouton: De Gruyter, 67-100.

Fletcher, William H. 1987. "Semantic factors in Dutch gender choice", in: William H. Fletcher (ed.), *Papers from the second interdisciplinary conference on Netherlandic studies*. Lanham, NY: University Press of America, 51-63.

Ford, L. J. 1998. A grammar of Emmi. [Unpublished Ph.D. dissertation, Australian National University.].

Frajzyngier, Zygmunt. 1993. *A grammar of Mupun*. Berlin: Reimer Verlag.

Franco, Ludovico, M. Rita Manzini and M. Leonardo Savoia. 2015. "N morphology and its interpretation: The neuter in Central Italian varieties and its implications", *Isogloss. Open Journal of Romance Linguistics*: 41-67.

Frankowsky, Maximilian and Dan Ke. 2016. "Humanness and classifiers in Mandarin Chinese", *Language and Cognitive Science* 2, 1: 55−67.

Friedrich, Paul. 1970. "Shape in grammar", *Language* 46, 2: 379–407.

Gao, Helena H. 2010. "A study of Swedish speakers' learning of Chinese noun classifiers", *Nordic Journal of Linguistics* 33, 2: 197-229.

Gao, Ming Y. and Barbara C. Malt. 2009. "Mental representation and cognitive conse-
quences of Chinese individual classifiers", *Language and Cognitive Processes* 24, 7-
8: 1124-1179.

Gao, Mingkai. 1948. *汉语语法论* [On Chinese grammar]. Beijing: Kaiming Press.

Gil, David 2013. "Numeral classifiers", in: Matthew S. Dryer and Martin Haspelmath (eds.),
The world atlas of language structures online. Leipzig: Max Planck Institute for
Evolutionary Anthropology. http://wals.info/chapter/55 (date of access: Mar 5 2022).

Greenberg, Joseph H. 1963. "Some universals of grammar with particular reference to the
order of meaningful elements", in: Joseph Greenberg (ed.), *Universals of Language*.
Cambridge, MA: MIT Press, 40-70.

Greenberg, Joseph H. 1972. "Numeral classifiers and substantival number: Problems in the
genesis of a linguistic type", *Working Papers in Language Universals* 9: 1-39.

Greenberg, Joseph H. 1974. "Studies in numerical system I: Double numeral systems",
Working Papers on Language Universals 14: 75-89.

Greenberg, Joseph H. 1978. "How does a language acquire gender markers?", in: Joseph H.
Greenberg, Charles A. Ferguson and Edith A. Moravcsik (eds.), *Universals of human
language. Vol. 3: Word structure*. Stanford, CA: Stanford University Press, 47-82.

Grinevald, Colette. 2000. "A morphosyntactic typology of classifiers", in: Gunter Senft
(ed.), *Systems of nominal classification*. Cambridge: Cambridge University Press,
51-92.

Grinevald, Colette. 2001. "Linguistics of classifiers", in: Neil J. Smelser and Paul B. Baltes
(eds.), *International encyclopedia of the social & behavioral sciences*. Amsterdam:
Elsevier, 1973-1978.

Grinevald, Colette. 2002. "Making sense of nominal classification systems", in: Ilse
Wischer and Gabriele Diewald (eds.), *New reflections on grammaticalization*. Am-
sterdam: John Benjamins, 259-275.

Grinevald, Colette. 2003. "Classifier systems in the context of a typology of nominal classi-
fication", in: Karen Emmorey (ed.), *Perspectives on classifier constructions in sign
language*. Mahwah, NJ: Lawrence Erlbaum, 91-109.

Grinevald, Colette. 2004. "Classifiers", in: Geert E. Booij, Christian Lehmann, Joachim
Mugdan and Stavros Skopeteas (eds.), *Morphologie/ Morphology; Ein Internation-
ales Handbuch zur Flexion und Wortbildung/An international handbook on inflec-
tion and word-formation*. Vol. 2. Berlin: Walter de Gruyter, 1016-1031.

Guan, Xiechu. 1953. *殷墟甲骨刻辞的语法研究* [A study of the grammar in the oracle inscription of Yinxu]. Beijing: Chinese Academy of Science.

Harvey, Mark and Nicholas Reid. 1997. *Nominal classification in Aboriginal Australia*. Amsterdam: John Benjamins.

He, Qun. 2011. "汉、维语个体量词"个"和 danɛ 的认知语义特征比较 [Cognitive semantic comparison between Chinese and Uyghur quantifier *ge* and *danɛ*]", *Journal of Minzu University of China (Philosophy and Social Sciences Edition)* 38, 3: 129-134.

Heath, Jeffrey. 1983. "Referential tracking in Nunggubuyu (Australia)", in: John Haiman and Pamela Munro (eds.), *Switch-reference and Universal Grammar: Proceedings of a Symposium on Switch Reference and Universal Grammar, Winnipeg, May 1981*. Amsterdam: John Benjamins, 129-149.

Her, One-Soon. 2012. "Distinguishing classifiers and measure words: A mathematical perspective and implications", *Lingua* 122, 14: 1668-1691.

Her, One-Soon. 2017. "Structure of numerals and classifiers in Chinese: Historical and typological perspectives and cross-linguistic implications", *Language and Linguisitics* 18, 1: 26–71.

Her, One-Soon, Ying-Chun Chen and Nai-Shing Yen. 2018. "Neural correlates of quantity processing of Chinese numeral classifiers", *Brain and Language* 176: 11-18.

Her, One-Soon and Yun-Ru Chen. 2013. Unification of numeral classifiers and plural markers: Empirical facts and implications. (Paper presented at the Proceedings of the 27th Pacific Asia Conference on Language, Information, and Computation (PACLIC 27), Taipei. (http://www.aclweb.org/anthology/Y13-1002).

Her, One-Soon and Chen-Tien Hsieh. 2010. "On the semantic distinction between classifiers and measure words in Chinese", *Language and Linguisitics* 11, 3: 527-551.

Her, One-Soon and Bing-Tsiong Li. 2019. "A single origin of numeral classifiers in Asia and Pacific: A hypothesis". (http://www3.nccu.edu.tw/~osh/Single%20Origin.pdf).

Her, One-Soon and Marc Tang. 2018. "A statistical explanation of the distribution of sortal classifiers in languages of the world via computational classifiers", *Journal of Quantitative Linguistics* 27, 2: 93-113.

Her, One-Soon and Hui-Chin Tsai. 2020. "Left is right, right is not: On the constituency of the classifier phrase in Chinese", *Language and Linguistics* 21, 1: 1-32.

Hewitt, John Napoleon Brinton. 1893. "Polysynthesis in the languages of the American Indians", *American Anthropologist* 6, 4: 381-408.

Hockett, Charles F. 1958. *A course in modern linguistics*. New York: Macmillan.

Hong, Cheng. 1961. "略论量词个的词源及其在唐以前的发展 [On the etymology of measure word ge and its development before the Tang]", *Journal of Nanjing University* (Nanjing) 2: 40-45.

Hsieh, Miao-Ling. 2008. *The internal structure of noun phrases in Chinese*. Taipei: Crane Publishing.

Hsu, Chun-Chieh, Shu-Hua Tsai, Chin-Lung Yang and Jenn-Yeu Chen. 2014. "Processing classifier–noun agreement in a long distance: An ERP study on Mandarin Chinese ", *Brain and language* 137: 14-28.

Hu, Qian 1993. The acquisition of Chinese classifiers by young Mandarin speaking children. [Unpublished Ph.D. dissertation, Boston University.].

Huang, C.-T. James and Masao Ochi. 2014. "Remarks on classifiers and nominal structure in East Asia", in: C.-T. James Huang and Feng-hsi Liu (eds.), *Peaches and plums*. (Language and Linguistics Monograph Series 54.) Taipei: Institute of Linguistics, Academia Sinica, 53-74.

Huang, Chu-Ren and Kathleen Ahrens. 2003. "Individuals, kinds and events: Classifier coercion of noun", *Language Sciences* 25, 4: 353-373.

Huang, Chu-Ren and Keh-jiann Chen. 2010. "Academia sinica balanced corpus of modern Chinese 4.0", *Academia Sinica.*, Retrieved Jan. 13, 2016. (http://asbc.iis.sinica.edu.tw/).

Huang, Shuping and Jenn-Yeu Chen. 2014. "The effects of numeral classifiers and taxonomic categories on Chinese and English speakers' recall of nouns", *Journal of East Asian Linguistics* 23: 27-42.

Imai, Mutsumi and Reiko Mazuka. 2007. "Language-relative construal of individuation constrained by universal ontology: Revisiting language universals and linguistic relativity", *Cognitive Science* 31, 3: 385-413.

Jackendoff, Ray S. 1977. *X-bar syntax: A study of phrase structure*. Cambridge, MA: MIT Press.

Jin, Fu-fen and Guo-hua Chen. 2002. "汉语量词的语法化 [On grammaticalization of Chinese classifiers]", *Journal of Tsinghua University* 17, 1: 8-14.

Jin, Guangjin, Hang Xiao and Li Fu. 2005. "现代汉语语料库建设及深加工 [Construction and deep processing of modern Chinese corpus]", *Applied Linguistics* 2: 111-120.

Jin, Jing. 2013. Functions of Chinese classifiers: A syntax-semantics interface account. [Unpublished Ph.D. dissertation, The Hong Kong Polytechnic University.].

Jin, Liling. 2018. "The processing of animacy in noun-classifier combinations in reading Korean: An ERP study", *Brain and Cognition* 126: 23-32.

Jones, Robert B. 1970. "Classifier constructions in Southeast Asia", *Journal of the American Oriental Society* 90, 1: 1-12.

Kemmerer, David. 2017. "Categories of object concepts across languages and brains: The relevance of nominal classification systems to cognitive neuroscience", *Language, Cognition and Neuroscience,* 32, 4: 401-424.

Kießling, Roland. 2018. "Niger-Congo numeral classifiers in a diachronic perspective", in: William McGregor and Søren K. Wichmann (eds.), *The diachrony of classification systems*. Amsterdam: John Benjamins, 33-76.

Kilarski, Marcin. 2013. *Nominal classification: A history of its study from the classical period to the present*. (Studies in the History of the Language Sciences 121.) Amsterdam: John Benjamins.

Kim, Christina. 2005. "Order and meaning: Numeral classifiers and specificity in Korean", in: John Alderete et al. (ed.), *Proceedings of the 24th West Coast Conference on Formal Linguistics.* Somerville, MA: Cascadilla Proceedings Project, 218-226.

Köpcke, Klaus-Michael, Klaus-Uwe Panther and David A. Zubin. 2010. "Motivating grammatical and conceptual gender agreement in German", in: Hans-Jörg Schmid and Susanne Handl (eds.), *Cognitive foundations of linguistic usage patterns*. Berlin: de Gruyter, 171-194.

Köpcke, Klaus-Michael and David A. Zubin. 1984. "Sechs Prinzipien für die Genuszuweisung im Deutschen: Ein Beitrag zur natürlichen Klassifikation", *Linguistische Berichte* 93: 26-50.

Lakoff, George. 1986. "Classifiers as a reflection of mind", in: Colette Craig (ed.), *Noun classes and categorization*. Amsterdam: John Benjamins.

LaPolla, Randy J. 1994. "Parallel grammaticalizations in Tibeto-Burman: Evidence of Sapir's 'Drift'", *Linguistics of the Tibeto-Burman Area* 17, 1: 61-80.

LaPolla, Randy J. and Chenglong Huang. 2003. *A grammar of Qiang with annotated texts and glossary*. Berlin: Walter de Gruyter.

Lee, Yunseok. 2014. *Classifiers in Korean*. Munich: Lincom.

Lehman, F. K. 1979. "Aspects of a formal theory of noun classifiers", *Studies in Language* 3, 2: 153-180.

Li, Jialei, Mingchen Sun and Jiajin Xu (eds.). 2022. *ToRCH2019 现代汉语平衡语料库* [ToRCH2019 modern Chinese balanced corpus]. China Research Center for Foreign Languages and Education, Beijing Foreign Studies University. (http://corpus.bfsu.edu.cn/info/1070/1335.htm).

Li, Jianping and Xiancheng Zhang. 2009. "泛指性量词 "枚/个" 的兴替及其动因——以出土文献为新材料 [The origin and its motivation of the general classifiers *mei* and *ge*: Based on new unearthed documents]", *Research in Ancient Chinese Language* 4: 64-72.

Li, Jinxi. [1924] 1992. *新著国语文法* [New Chinese grammar]. Beijing: Commercial Press. [Originally published, Beijing: Commercial Press.].

Li, Peggy, David Barner and Becky H. Huang. 2008. "Classifiers as count syntax: Individuation and measurement in the acquisition of Mandarin Chinese", *Language Learning and Development* 4, 4: 1-42.

Li, Wendan. 2000. "The pragmatic function of numeral-classifiers in Mandarin Chinese", *Journal of Pragmatics* 32, 8: 1113-1133.

Li, XuPing. 2011. On the semantics of classifiers in Chinese. [Unpublished Ph.D. dissertation, Bar Ilan University.].

Li, XuPing. 2013. *Numeral classifiers in Chinese: The syntax-semantics interface*. Berlin: De Gruyter Mouton.

Li, XuPing and Walter Bisang. 2012. "Classifiers in Sinitic languages: From individuation to definiteness-marking", *Lingua* 122: 335-355.

Li, Yanzhi and Yicheng Wu. 2018. "The definite interpretation of [CL+ N] expressions in Sinitic languages: Synchronic and diachronic perspectives", *Taiwan Journal of Linguistics* 16, 2: 115-157.

Li, Yen-hui Audrey. 1999. "Plurality in a classifier language", *Journal of East Asian Linguistics* 8, 1: 75-99.

Lichtenberk, Frantisek. 1983. "Relational classifiers", *Lingua* 60, 2-3: 147-176.

Liu, Cixin. [2008] 2014. *三体.* [The three-body problem]. (Translated by Ken Liu.) New York: Tor Books. [Originally published, Chongqing: Chongqing Publishing House.].

Liu, Jianmin. 2016. "小议个体量词"累"与"果" [The origin of individual measure word *lei* and *guo*]", *Journal of Linguistic Sciences* 15, 03: 264-267.

Liu, Shiru. 1965. *魏晋南北朝量词研究* [On quantifiers in Weijin and the Northern and Southern Periods]. Beijing: Zhonghua Book Company.

Loke, Kit-Ken. 1997. "The grammaticalization and regrammaticalization of Chinese numeral classifier morphemes", *Journal of Chinese Linguistics* 25: 1-20.

Lucy, John A. 1992. *Grammatical categories and cognition*. Cambridge: Cambridge University Press.

Lucy, John A. 2000. "Systems of nominal classification: A concluding discussion", in: Gunter Senft (ed.), *Systems of nominal classification*. Cambridge: Cambridge University Press, 326-341.

Luraghi, Silvia. 2011. "The origin of the Proto-Indo-European gender system: Typological considerations", *Folia Linguistica* 45, 2: 435-463.

Lyons, John. 1968. *Introduction to theoretical linguistics*. Cambridge: Cambridge University Press.

Lyons, John. 1977. *Semantics*. Vol. 2. Cambridge: Cambridge University Press.

Ma, Aimin. 2011. "《汉语大字典》、《汉语大词典》个体量词编纂疏误 [Mistakes in the compilation of individual quantifiers in *The Big Dictionary of Chinese Word* and *The Big Dictionary of Chinese*]", *Journal of Nanchang University* 42, 1: 148-151.

Ma, Aimin. 2015. *汉语个体量词的产生与发展* [The origin and development of Chinese individual classifiers]. Beijing: China Social Science Press.

Ma, Jianzhong. [1898] 2010. *马氏文通* [Mr Ma's grammar]. Shanghai: Shanghai Commercial Press. [Originally published, Beijing: Commercial Press.].

Ma, Xiaoyi. 1999. "Parallel text collections at linguistic data consortium", *Proceedings of Machine Translation Summit VII*. Singapore, 346-348.

Marnita, Rina. 1996. Classifiers in Minangkabau. [Unpublished MA thesis, Australian National University, Canberra.].

Martins, S. A. 1994. Análise da morfosintaxe da língua Dâw (Makú-Kamã) e sua classificação tipológica. [Unpublished MA thesis, Universidade Federal de Santa Catarina, Florianópolis.].

Mathieu, Eric. 2012. "On the mass/count distinction in Ojibwe", in: Diane Massam (ed.), *Count and Mass across Languages*. Oxford: Oxford University Press, 172-198.

Matthews, Peter Hugoe. 1997. *The concise Oxford dictionary of linguistics*. Oxford: Oxford University Press.

McEnery, Tony and Richard Xiao (eds.). 2004. *The Lancaster corpus of mandarin chinese*. Lancaster: Lancaster University. (https://web.ar-chive.org/web/20200508000520id_/https://eprints.lancs.ac.uk/id/eprint/71/1/The_Lancaster_Corpus_of_Mandarin_Chinese.pdf).

McGregor, William B. and Søren Wichmann (eds.). 2018. *The diachrony of classification systems*. Amsterdam: John Benjamins.

Meng, Fanjie and Rulong Li. 2011. "量词"片"的语法化 [On grammaticalization of Chinese measure word *pian*(片)]", *Studies in Language and Linguistics* 31, 03: 66-70.

Messineo, Cristina and Paola Cúneo. 2019. "Multifunctionality of deictic classifiers in the Toba language (Guaycuruan)", in: Alexandra Y. Aikhenvald and Elena I. Mihas (eds.), *Genders and Classifiers*. Oxford: Oxford University Press, 197-221.

Mithun, Mariann. 1986. "The convergence of noun classification systems", in: Colette Craig (ed.), *Noun classes and categoriztion*. Amsterdam: John Benjamins, 379-397.

Mithun, Marianne. 2014. "Gender and culture", in: Greville G. Corbett (ed.), *The expression of gender*. Berlin: De Gruyter Mouton, 131-160.

Morrison, Robert. 1815. *A grammar of the Chinese language*. Serampore: Mission Press.

Niu, Baoyi. 2010. "心智模拟概念化——量词"把"的语义概念化研究 [Conceptualization through mental simulation: On the conceptualization of the measure *ba*]", *Foreign Language Education* 31, 5: 1-5, 14.

Parker, Enid M. and Richard J. Hayward. 1985. *An Afar-English-French dictionary (with grammatical notes in English)*. London: School of Oriental and African Studies, University of London.

Passer, Matthias Benjamin. 2016. *The typology and diachrony of nominal classification*. Utrecht: LOT.

Pawley, Andrew. 2002. "Using 'he' and 'she' for inanimate referents in English: Questions of grammar and world view", in: Nick J. Enfield (ed.), *Ethnosyntax: Explorations in grammar and culture*. Oxford: Oxford University Press, 110-137.

Pet, Willem J. A. 2011. *A grammar sketch and lexicon of Arawak (Lokono Dian)*. Dallas, Texas: SIL International.

Peyraube, Alain. 1991. "Some remarks on the history of Chinese classifiers", *Santa Barbara Papers in Linguistics* 3: 106-126.

Plaster, Keith and Maria Polinsky. 2007. "Women are not dangerous things: Gender and categorization", *Harvard Working Papers in Linguistics 12*.

Rankin, Robert L. 2004. "The history and development of Siouan positionals with special attention to polygrammaticalization in Dhegiha", *STUF - Language Typology and Universals* 57, 2/3: 202–228.

Reid, Nicholas. 1997. "Class and classifier in Ngan'gityemerri", in: Mark Harvey and Nicholas Reid (eds.), *Nominal classification in Aboriginal Australia*. Amsterdam: John Benjamins, 165-228.

Ren, Liang'e, Hongren Sun and Jianding Yang (eds.). 2010. *Chinese-English Parallel Corpus of A Dream of Red Mansions*. (http://corpus.usx.edu.cn/) (date of access: Apr 9 2022).

Rice, Curt. 2006. "Optimizing gender", *Lingua* 116, 9: 1394-1417.

Rodrigues, Aryon D. 2012. "Nominal classification in Kariri", *Revista Brasileira de Linguística Antropológica* 4, 2: 257-265.

Royen, Gerlach. 1929. *Die nominalen Klassifikations-Systeme in den Sprachen der Erde: Historisch-kritische Studie, mit besonderer Berücksichtigung des Indogermanischen.* (Anthropos Linguistische Bibliothek 4.) Vienna: Mechitharisten-Buchdruckerei.

Saalbach, Henrik and Mutsumi Imai. 2012. "The relation between linguistic categories and cognition: The case of numeral classifiers", *Language and Cognitive Processes* 27, 3: 381-428.

Salehuddin, Khazriyati, Heather Winskel and Marlyna Maros. 2011. "The pragmatic functions of numeral classifiers In Modern Malay written corpus", *GEMA Online Journal of Language Studies* 11, 2: 137-153.

Sanches, Mary and Linda Slobin. 1973. "Number classifiers and the plural marking: An implicational universal", *Working Papers in Language Universals* 11: 1-22.

Sandman, Erika and Francesca Di Garbo. In press. "Contact-induced reduction, loss, and emergence of numeral classifiers: Two case studies from East Asia", in: Marc Allassonnière-Tang and Marcin Kilarski (eds.), *Nominal classification in Asia: Functional and diachronic perspectives*. (Current Issues in Linguistic Theory.) Amsterdam: John Benjamins.

Sands, Kristina. 1995. "Nominal classification in Australia", *Anthropological Linguistics* 37: 247-346.

Seifart, Frank. 2010. "Nominal classification", *Language and Linguistics Compass* 4, 8: 719-736.

Seifart, Frank. 2018. "The semantic reduction of the noun universe and the diachrony of nominal classification", in: William B. McGregor and Søren Wichmann (eds.), *The diachrony of classification systems*. Amsterdam: John Benjamins, 9-32.

Seiler, Walter. 1986. "Noun-classificatory verbal prefixes as reanalysed serial verbs", *Lingua* 68, 2/3: 189-207.

Selkirk, Elisabeth. 1977. "Some remarks on noun phrase structure", in: Peter W. Culicover, Thomas Wasow and Adrian Akmajian (eds.), *Formal syntax*. New York: Academic Press, 285–316.

Senft, Gunter. 2007. "Nominal classification", in: Dirk Geeraerts and Hubert Cuyckens (eds.), *The Oxford handbook of cognitive linguistics*. Oxford: Oxford University Press, 676-696.

Senft, Gunter. 2012. "Nominal classification", in: Dirk Geeraerts and Hubert Cuyckens (eds.), *The Oxford handbook of cognitive linguistics*. (DOI: 10.1093/oxfordhb/9780199738632.013.0026) (date of access: 30 Oct. 2021).

Senft, Gunter (ed.). 2000. *Systems of nominal classification*. Cambridge: Cambridge University Press.

Serzisko, Fritz. 1982. "Gender, noun class and numeral classification: A scale of classificatory techniques", in: René Dirven and Günter Radden (eds.), *Issues in the theory of universal grammar*. Tübingen: Gunter Narr, 95-123.

Shao, Tiansong. 2015. "黑水城出土宋代汉文社会文献中的个体量词 [On the individual classifiers in the Song Dynasty Chinese documents excavated in Heishui City]", *Journal of School of Chinese Language and Culture Nanjing Normal University* 3: 166-173.

Shaul, David L. 1986. *Topics in Nevome syntax*. Berkeley, CA: University of California Press.

Simpson, Andrew. 2005. "Classifiers and DP structure in Southeast Asia", in: Guglielmo Cinque and Richard Kayne (eds.), *The Oxford handbook of comparative syntax*. Oxford: Oxford University Press, 1-22.

Sinclair, John McHardy (ed.). 2004. *Trust the text: Language, corpus and discourse*. London: Routledge.

Sinnemäki, Kaius. 2019. "On the distribution and complexity of gender and numeral classifiers", in: Francesca Di Garbo and Bernhard Wälchli (eds.), *Grammatical gender and linguistic complexity*. Berlin: Language Science Press, 134-200.

Song, Jiang. 2017. *The semantics of Chinese classifiers and linguistic relativity*. London: Routledge.

Song, Ju. 2014. "《马氏文通》前西人的汉语量词研究——以《语言自迩集》为核心 [The Early Westerner's Study of Chinese Measure Word before Ma Shi Wen Tong: Based on Yü-yen Tzŭ-êrh Chi]", *Studies in Language and Linguistics* 4: 93-101.

Srinivasan, Mahesh. 2010. "Do classifiers predict differences in cognitive processing? A study of nominal classification in Mandarin Chinese", *Language and Cognition* 2, 2: 177-190.

Steele, Susan. 1978. "Word order variation: A typological study", in: Joseph H. Greenberg, Charles A. Ferguson and Edith A. Moravcsik (eds.), *Universals of Human Language*. Vol. 4: Syntax. Stanford, CA: Stanford University Press, 585-623.

T'sou, Benjamin K. 1976. "The structure of nominal classifier systems", in: Philip N. Jenner, Laurence C. Thompson and Stanley Starosta (eds.), *Austroasiatic studies*. San Diego: University of Hawaii Press, 1215-1247.

Tai, James H-Y. 1992. "Variation in classifier systems across Chinese dialects", *Chinese Languages and Linguisitics* 1: 587-608.

Tai, James H-Y. 1994. "Chinese classifier systems and human categorization", in: Mattew Chen and Ovid Tzeng (eds.), *In honor of William S.-Y. Wang: Interdisciplinary studies on language and language change.* Taipei: Pyramid Press, 1-17.

Tai, James and Lianqing Wang. 1990. "A semantic study of the classifier tiao (条)", *Journal of the Chinese Language Teachers Association* 25, 1: 35-56.

Tang, Chih-Chen Jane. 1990. Chinese phrase structure and the extended X'-theory. [Unpublished Ph.D. dissertation, Cornell University.].

Tang, Chih-Chen Jane. 2005. "Nouns or classifiers: A non-movement analysis of classifiers in Chinese", *Language and Linguistics* 6, 3: 431-472.

Tang, Marc, Ying-Chun Chen, Nai-Shing Yen and One-Soon Her. 2021. "Investigating the branching of Chinese classifier phrases: evidene from speech perception and production", *Journal of Chinese Linguistics* 49, 1: 71-105.

Tang, Marc and One-Soon Her. 2019. "Insights on the Greenberg-Sanches-Slobin generalization: Quantitative typological data on classifiers and plural markers", *Folia Linguistica* 53, 2: 297-331.

Tucker, G. Richard, Wallace E. Lambert and André Rigault. 1977. *The French speaker's skill with grammatical gender: An example of rule-governed behavior*. The Hague: Mouton.

Wagner, Susanne. 2003. Gender in English pronouns: Myth and reality. [Unpublished Ph.D. dissertation, University of Freiburg.].

Wang, Chongyang. 2018. ""个化"问题及量词"个"表量的局限性分析 [A brief analysis of the problem of "个化" and the semantic limitations of quantifier "个"]", *Journal of Mudanjiang Normal University* 04: 101-105.

Wang, Li. [1954] 1985. *中国现代语法* [Modern Chinese grammar]. Beijing: Commercial Press. [Originally published, Beijing: China Youth Press.].

Wang, Li. [1956] 2004. *汉语史稿* [A preliminary history of Chinese]. Beijing: Zhonghua Book Company. [Originally published, Beijing: Zhonghua Book Company.].

Wang, Lianqing. 1994. Origin and development of classifiers in Chinese. [Unpublished Ph.D. dissertation, Ohio State University, Columbus, OH.].

Wang, Sahoxin. 2010. "汉语史上名量词语法化问题 [On grammaticalization of noun quantifiers in the history of the Chinese language]", *Journal of Shananxi Normal University (Philosophy and Social Sciences Edition)* 39, 3: 168-176.

Wang, Xiaolu and Wan Ren. 2017. "词义联想与汉语二语量词习得 [Lexical association and CSL learners' acquisition of Chinese numeral classifiers]", *Chinese as a Second Language Research* 6, 1: 129–147.

Welmers, W. E. 1973. *African language structures*. Berkley, CA: University of California Press.

Wu, Jiun-Shiung and One-Soon Her. 2021. "Taxonomy of numeral classifiers: A formal semantic proposal", in: Chungmin Lee, Young-Wha Kim and Byeong-uk Yi (eds.), *Numeral classifiers and classifier languages: Chinese, Japanese, and Korean*. New York: Routledge, 40-71.

Wu, Yayun. 2014a. "汉语个体量词" 数+量+名" 结构的历时形成过程 [The diachrony of the construcion [NUM+CL+N] of Chinese numeral classifiers]", *Chinese Linguistics* 47, 3: 83-90.

Wu, Yayun. 2014b. *汉语个体量词的功能及其历时形成过程* [The functions of classifiers in Mandarin Chinese: A diachronic perspective]. [Unpublished Ph.D. dissertation, National University of Singapore, Department of Chinese Studies, Singapore.].

Xiao, Hang. 2010. "现代汉语通用平衡语料库建设与应用 [Construction and application of modern Chinese Balanced Corpus]", *Journal of Chinese Language Teaching* 106.

Xiao, Hang. 2016. *语料库词义标注研究* [A study of semantic annotation in corpus]. Kunming: Yunnan Education Publishing House.

Xiao, Richard (ed.). 2004. *The Babel English-Chinese Parallel Corpus*. UCREL, Lancaster. (http://www.lancs.ac.uk/fass/projects/corpus/babel/babel.htm).

Xiao, Richard and Tony McEnery. 2010. *Corpus-based contrastive studies of English and Chinese*. New York: Routledge.

Xu, Jiajin (ed.). 2019a. *CLIPS corpus*. (http://corpus.bfsu.edu.cn/info/1070/1335.htm).

Xu, Jiajin (ed.). 2019b. *TED 英汉平行演讲语料库* [TED English Chinese parallel corpus of speech]. (http://corpus.bfsu.edu.cn/info/1070/1556.htm).

Xu, Jiajin, Maocheng Liang and Yunlong Jia. 2012. BFSU ParaConc 1.2.1. National Research Center for Foreign Language Education, Beijing Foreign Studies University.

Xu, Xiaodong, Xiaoming Jiang and Xiaolin Zhou. 2013. "Processing biological gender and number information during Chinese pronoun resolution: ERP evidence for functional differentiation", *Brain and Cognition* 81, 2: 223-236.

Xun, Endong, Gaoqi Rao, Xiaoyue Xiao and Jiaojiao Zang. 2016. "The construction of the BCC Corpus in the age of big data", *Corpus Linguistics and Linguistic Theory* 3, 1: 93-109, 118.

Yamamoto, Kasumi and Frank Keil. 2000. "The acquisition of Japanese numeral classifiers: Linkage between grammatical forms and conceptual categories", *Journal of East Asian Linguistics* 9, 4: 379-409.

Yang, Linwei. 2021. "TED 演讲英汉平行语料库 [TED English Chinese parallel corpus of speeches]". (http://icorpus.net/application/ted/).

Zhan, Weidong, Rui Guo, Baobao Chang, Yirong Chen and Long Chen. 2019. "The building of the CCL corpus: Its design and implementation", *Corpus Linguistics* 6, 1: 71-86.

Zhan, Weidong, Rui Guo and Yirong Chen (eds.). 2003. *The CCL Corpus of Chinese Texts: 700 million Chinese Characters, the 11th Century B.C. - present*. (http://ccl.pku.edu.cn:8080/ccl_corpus).

Zhang, Hong. 2007. "Numeral classifiers in Mandarin Chinese", *Journal of East Asian Linguistics* 16, 1: 43-59.

Zhang, Hong, Jianding Yang, Hongren Sun, Shibao Zhou and Shirong Liu (eds.). 2011. *E-C Parallel Corpus of Moment in Peking*. (http://corpus.usx.edu.cn/) (date of access: Apr 9 2022).

Zhang, Jie and Xiaofei Lu. 2013. "Variability in Chinese as a foreign language learners' development of the Chinese numeral classifier system", *The Modern Language Journal* 97: 46-60.

Zhang, Jinqiao and Xiao Xiao. 2012. "中级水平留学生言语产生中汉语个体量词的通达机制 [On the access mechanism of Chinese individual measure words in the language production of overseas students at intermediate level]", *TCSOL Studies* 02: 23-29.

Zhang, Niina Ning. 2011. "The constituency of classifier constructions in Mandarin Chinese", *Taiwan Journal of Linguistics* 9, 1: 1-50.

Zhang, Niina Ning. 2012. *Numeral classifier structures in Mandarin Chinese*. Berlin: De Gruyter.

Zhang, Sihong. 2013. A reference grammar of Ersu, a Tibeto-Burman language from China. [Unpublished Ph.D. dissertation, James Cook University.].

Zhou, Shao. 2014. "量词"个"与名词的组合倾向及中性选择机制 [On construction rule of classifier "ge" and nouns and its neutral-selective mechanism]", *TCSOL Studies* 1: 86-93.

Zubin, David A and Mitsuaki Shimojo. 1993. "How 'general' are general classifiers? With special reference to *ko* and *tsu* in Japanese", in: Joshua S. Guenter, Barbara A. Kaiser and Cheryl C. Zoll (eds.), *Proceedings of the Nineteenth Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Semantic Typology and Semantic Universals*. Berkeley, CA: Berkeley Linguistics Society, 490-502.

Zubin, David A. and Klaus-Michael Köpcke. 1984. "Affect classification in the German gender system", *Lingua* 63: 41-96.

Zubin, David A. and Klaus-Michael Köpcke. 2009. "Gender control: Lexical or conceptual", in: Patrick O. Steinkrüger and Manfred Krifka (eds.), *On inflection*. Berlin: Mouton de Gruyter, 237-262.

Appendix 1: The timeline of Ancient Chinese history[26]

Years	Dynasties or Periods of coexisting Kingdoms	
c. 100, 0000-2183 BC	Pre-history of Ancient China	
c. 2183-1600 BC	Xia Dynasty	
c. 1600-1050 BC	Shang Dynasty	
c. 1050-771 BC	Western Zhou Dynasty	
770-256 BC	Eastern Zhou Dynasty	
770-221 BC	Pre-Qin Period	Spring and Autumn period (770-403 BC)
		Warring State period (403-221 BC)
221-206 BC	Qin Dynasty	
206 BC-9	Han Dynasty	Western Han Dynasty
9-25		Period of Xin Mang
25-220		Eastern Han Dynasty
220-265	Wei-Jin Periods	Kingdom of Wei
		Kingdom of Shu
		Kingdom of Wu
265-316		Western Jin
317-420		Eastern Jin
304-439		Sixteen States
420-589	Northern and Southern Dynasties	Northern Zhou (557-581)
		Southern Dynasties (420-589)
581-618	Sui Dynasty	
618-907	Tang Dynasty	
907-960	Five Dynasties and Ten Kingdoms Period	Later Liang (907-923)
		Later Tang (923-936)
		Later Jin (936-946)
		Later Han (947-950)
		Later Zhou (951-960)
		Liao (907-1126)
960-1276	Song Dynasty	Northern Song (960-1127)
		Southern Song (1127-1276)
		Great Jin (1115-1234)
1279-1368	Yuan Dynasty	
1368-1644	Ming Dynasty	
1644-1911	Qing Dynasty	

---

[26] Adapted based on Ebrey (1996: 338-340)

Appendix 2: List of Chinese numeral classifiers

Types	No.	Numeral classifiers		Semantics	Frequency in Corpus 1	Frequency in Corpus 2
Entity	1	个	*gè* 'CLF:GENERAL'	general	46, 43%	0,76%
	2	位	*wèi* 'CLF:INDIVIDUAL, RESPECT'	human: respect	4, 16%	
	3	名	*míng* 'CLF:IDENTITY'	human: identity	2, 96%	
	4	只	*zhī* 'CLF:SINGLE'	nonhumanness	2, 49%	
	5	条	*tiáo* 'CLF:SLENDER'	shape: one-dimension	2, 27%	5.35%
	6	家	*jiā* 'CLF:HOUSEHOLD'	organization	2, 13%	
	7	项	*xiàng* 'CLF:ITEM, PRO-JECT'	unspecified	1, 99%	
	8	块	*kuài* 'CLF:LUMP-LIKE'	shape: three-dimension	1, 27%	52.77%
	9	张	*zhāng* 'CLF:SPREAD-ING.OPEN/FLAT'	shape: two-dimension	1, 13%	2,10%
	10	座	*zuò* 'CLF:SEAT, PEDES-TAL, BASE'	salient physical feature	0, 97%	0,19%
	11	根	*gēn* 'CLF:ROOT, STICK-SHAPE'	shape: two-dimension	0, 84%	
	12	部	*bù* 'CLF:DEMO'	salient physical feature	0, 72%	0,19%
	13	本	*běn* 'CLF:BOOK'	salient physical feature	0, 67%	0,19%
	14	层	*céng* 'CLF:LAYER'	shape: two-dimension	0, 67%	4.40%
	15	间	*jiān* 'CLF:ROOM'	salient physical feature	0, 63%	0,76%
	16	辆	*liàng* 'CLF:VEHICLE, CAR'	salient physical feature	0, 63%	
	17	颗	*kē* 'CLF:ROUNDISH'	shape: three-dimension, size	0, 61%	0,96%
	18	支	*zhī* 'CLF:BRANCH'	shape: two-dimension	0, 58%	
	19	片	*piàn* 'CLF:FLAT/THIN'	shape: two-dimension	0, 57%	17.21%
	20	篇	*piān* 'CLF:ARTICLE'	salient physical feature	0, 54%	0,38%
	21	把	*bǎ* 'CLF:HANDLE'	salient physical feature	0, 49%	
	22	台	*tài* 'CLF:PLATFORM, MECHANISM'	salient physical feature	0, 48%	

23	棵	*kē* 'CLF:PLANT'	salient physical feature	0, 46%	0,19%
24	所	*suō* 'CLF:PLACE'	salient physical feature	0, 43%	
25	首	*shǒu* 'CLF:SONG, POEM'	salient physical feature	0, 40%	
26	幅	*fú* 'CLF:WIDTH, PICTURE'	shape: two-dimension	0, 31%	0,57%
27	处	*chù* 'CLF:LOCATION'	salient physical feature	0, 30%	
28	封	*fēng* 'CLF:SEALING, ENVELOP'	salient physical feature	0, 30%	
29	股	*gǔ* 'CLF:STRAND'	shape: two-dimension	0, 28%	0,38%
30	架	*jià* 'CLF:FRAMEWORK'	shape: three-dimension	0, 25%	
31	枚	*méi* 'CLF:GENERAL, ROUND.PIECE'	shape: three-dimension	0, 22%	
32	艘	*sōu* 'CLF:SHIP'	salient physical feature	0, 22%	
33	门	*mén* 'CLF:GATE, BRANCH'	salient physical feature	0, 21%	
34	段	*duàn* 'CLF:SEGMENT'	shape: two-dimension	0, 19%	4,40%
35	句	*jù* 'CLF:SENTENCE'	salient physical feature	0, 19%	0,57%
36	扇	*shàn* 'CLF:FAN'	shape: two-dimension	0, 18%	0,19%
37	具	*jù* 'CLF:LONG.AND.STIFF'	shape: two-dimension	0, 16%	
38	朵	*duǒ* 'CLF:FLOWER-LIKE'	shape: three-dimension	0, 16%	0,57%
39	份	*fèn* 'CLF:SHARE/PORTION'	unspecified	0, 16%	
40	头	*tóu* 'CLF:HEAD'	animal: salient physical feature	0, 15%	
41	副	*fù* 'CLF:SET, ATTITUDE'	unspecified	0, 13%	
42	粒	*lì* 'CLF:GRAIN-LIKE'	shape: three-dimension, size	0, 13%	0,57%
43	道	*dào* 'CLF:COURSE-LIKE, PATH-LIKE'	shape: two-dimension	0, 12%	
44	则	*zé* 'CLF:CLAUSE, ENTRY'	unspecified	0, 09%	
45	株	*zhū* 'CLF:STALK'	plant: salient physical feature	0, 10%	
46	匹	*pǐ* 'CLF:HORSE, HORSE-LIKE.ANIMAL'	animal: salient physical feature	0, 09%	
47	盏	*zhǎn* 'CLF:DISH'	dish: salient physical feature	0, 09%	0,19%
48	团	*tuán* 'CLF:BALL'	shape: three-dimension	0, 07%	
49	滴	*dī* 'CLF:DROP'	shape: three-dimension, size	0, 07%	1,72%
50	堵	*dǔ* 'CLF:BLOCK'	shape: two-dimension	0, 07%	
51	顶	*dǐng* 'CLF:TOP'	salient physical feature	0, 06%	
52	栋	*dòng* 'CLF:BUILDING, BEAM'	salient physical feature	0, 06%	
53	柄	*bǐng* 'CLF:STIPE'	salient physical feature	0, 04%	
54	代	*dài* 'CLF:GENERATION'	generation	0, 04%	
55	枝	*zhī* 'CLF:BRANCH'	shape: two-dimension	0, 04%	0,19%

	56	泓	*hóng* 'CLF:DEEP.AND.VAST'	shape: two-dimension	0, 03%	
	57	撇	*piě* 'CLF:LEFT.FALLING'	shape: two-dimension	0, 03%	0,19%
	58	杆	*gān* 'CLF:STEM'	shape: two-dimension	0, 03%	
	59	口	*kǒu* 'CLF:OPENING'	salient physical feature	0, 03%	0,19%
	60	尾	*wěi* 'CLF:TAIL'	animal: salient physical feature	0, 03%	
	61	幢	*zhuàng* 'CLF:BUILDING, PILLAR'	salient physical feature	0, 03%	
	62	点	*diǎn* '*CLF:POINT*'	shape: three-dimension	0, 01%	
	63	方	*fāng* 'CLF:SQUARE'	shape: two-dimension	0, 01%	
	64	例	*lì* 'CLF:CASE'	unspecified	0, 01%	
	65	缕	*lǔ* 'CLF:WISP'	shape: two-dimension	0, 01%	0.19%
	66	床	*chuáng* 'CLF:BED'	salient physical feature	0, 01%	
	67	剂	*jì* 'CLF:DOSE'	salient physical feature	0, 01%	0,19%
	68	节	*jié* 'CLF:SECTION'	shape: two-dimension	0, 01%	
	69	截	*jié* 'CLF:CUTTING'	shape: two-dimension	0, 01%	1,15%
	70	列	*liè* 'CLF:ROW'	shape: two-dimension	0, 01%	
	71	帘	*lián* 'CLF:CURTAIN'	salient physical feature	0, 01%	
	72	面	*miàn* 'CLF:SURFACE'	shape: two-dimension	0, 01%	
	73	腔	*qiāng* 'CLF:TONE, TUNE'	salient physical feature	0, 01%	
	74	滩	*tān* 'CLF:POOL.LIKE'	shape: two-dimension	0, 01%	
	75	弯	*wān* 'CLF:CURVE-LIKE'	shape: three-dimension	0, 01%	
	76	湾	*wān* 'CLF:BEND'	shape: three-dimension	0, 01%	
	77	页	*yè* 'CLF:PAGE'	salient physical feature	0, 01%	0,38%
	78	尊	*zūn* 'CLF:RESPECT'	salient physical feature	0, 01%	
	79	绺	*liǔ* 'CLF:TUFT, LOCK'	salient physical feature		0,38%
Entity/ Event	80	件	*jiàn* 'CLF:PIECE'	unspecified: piece	1, 15%	0,38%
	81	轮	*lún* 'CLF:WHEEL, ROUND'	shape: three-dimension, event	0, 09%	
	82	圈	*quān* 'CLF:CIRCLE'	shape: three-dimension, event	0, 06%	
	83	盘	*pán* 'CLF:PLATE, GAME'	shape: two-dimension event: game	0, 04%	
Event	84	场	*cháng* 'CLF:VENUE'	event: venue	1, 93%	0,19%
	85	次	*cì* 'CLF:TIME'	event: time	1, 49%	
	86	声	*shēng* 'CLF:SOUND'	event: sound	0, 33%	0,19%
	87	顿	*dùn* 'CLF:SPELL, SES-SION'	event: short time	0, 30%	0,38%
	88	笔	*bǐ* 'CLF:PEN'	event: transaction	0, 19%	0,19%
	89	出	*chū* 'CLF:EXIT, EN-TRANCE'	event: duration of time	0, 15%	

	90	阵	*zhèn* 'CLF:PEIROD'	sporadic event	0, 13%	
	91	记	*jì* 'CLF:STRIKE'	event: duration of time	0, 10%	
	92	桩	*zhuāng* 'CLF:STAKE'	event: piece	0, 07%	
	93	届	*jiè* 'CLF:SESSION'	event: due time	0, 03%	
	94	幕	*mù* 'CLF:SCENE'	event: act of a play	0, 03%	
	95	泡	*pào* 'CLF:BUBBLE, BREWING'	event: procedure	0, 03%	
	96	堂	*táng* 'CLF:HALL'	event: class	0, 03%	
	97	宗	*zōng* 'CLF:FACTION'	event: trade/transaction	0, 03%	
	98	餐	*cān* 'CLF:DISH, MEAL'	event: meal	0, 01%	
	99	番	*fān* 'CLF:COURSE'	event: duration of time	0, 01%	
	100	趟	*tàng* 'CLF:TRIP'	event: journey	0, 01%	
	101	席	*xí* 'CLF:MAT'	event: talk, feast	0, 01%	
Kind	102	种	*zhǒng* 'CLF:KIND, GENERAL'	kind: general	13, 49%	
	103	类	*lèi* 'CLF:CATEGORY'	kind: category	0, 39%	1,15%
	104	款	*kuǎn* 'CLF:STYLE'	kind: style	0, 33%	
	105	样	*yàng* 'CLF:SAMPLE'	kind: sample	0, 03%	
Total					6700	523