

Quantifying Constructions in English and Chinese

A Corpus-Based Contrastive Study

Tony McEnery¹ and Richard Xiao¹

Abstract

Quantifiers are a linguistic concept that mirrors quantity in reality. They indicate ‘how many’ or ‘how much’, for example, the number of entities denoted by a noun, the count of actions or events, the length of time, and the distance in space. All human languages have linguistic devices that express such ideas, though the encoding of natural language semantics can vary from language to language. This paper compares quantifying constructions in English and Chinese on the basis of comparable corpora of spoken and written data in the two languages. We will focus on classifiers in Chinese and their counterparts in English, as well as the interaction between quantifying constructions and progressives, which is normally ruled out by aspect theory, with the aim of addressing the following research questions:

- What linguistic devices are used in Chinese and English for quantification?
- How different (or similar) are classifiers in Chinese as a classifier language and in English as a non-classifier language?
- Can quantifiers interact with progressives in English and Chinese if such interactions are theoretically ruled out by aspect theory?

Before these research questions are explored in detail, it is appropriate to first present the principal data used in this study, which includes two written corpora and two spoken corpora. The Freiburg-LOB (FLOB) corpus is a recent update of LOB, which is composed of approximately one million tokens of written British English sampled proportionally from fifteen text categories published in the early 1990s (Hundt et al. 1998). The Lancaster Corpus of Mandarin Chinese (LCMC) was designed as a Chinese match for FLOB and created using the same sampling criteria, representing written Mandarin Chinese published in China in the corresponding sampling period (McEnery et al. 2003). The two spoken corpora are BNCdemo and CallHome Mandarin. BNCdemo is the demographically sampled component of the British National Corpus (BNC), which contains four million tokens of transcripts of conversations recorded around the early 1990s. The CallHome Mandarin Transcripts corpus, which was released by the LDC, comprises 120 transcripts of 5-to-10-minute telephone conversations recorded in the first half of the 1990s between native Chinese speakers living overseas and their families in China, amounting to approximately 300,000 tokens. While telephone calls differ from face-to-face conversations alongside some dimensions (Biber 1988), the sampling periods of two spoken corpora are roughly comparable. A practical reason for using the CallHome corpus is that this dataset is closest to BNCdemo which is available to us.

¹ Department of Linguistics and English Language, Lancaster University
e-mail: a.mcenery@lancaster.ac.uk, z.xiao@lancaster.ac.uk

In the remaining sections of this article, we will first explore classifiers in Chinese and English, on the basis of which the two will be compared. We will then discuss the interaction of the progressive with quantifying constructions in the two languages.

1. Classifiers in Chinese

Chinese is generally recognised as a ‘classifier language’ not only because of its large inventory of classifiers but also because the use of classifier is mandatory in this language. While it is difficult to give an exact number of classifiers because of the fuzzy boundaries between classifiers and nouns on the one hand and between classifiers and numerals on the other hand (cf. Xing 1993; Li 2000), it has been estimated that there are 500–600 commonly used classifiers in Chinese (cf. Guo 1987: 10). In the LCMC and CallHome corpora used in this study, a total of 421 types of classifiers are found (or 342 types when reduplicated classifiers and those with a suffix 儿 *er* or 子 *zi* are collapsed into their stems). In terms of tokens, classifiers account for 2.48% of LCMC and 3.13% of CallHome. Chinese does not make a distinction between count and noncount nouns. A classifier is obligatory when a noun is quantified in modern Chinese, whether its English equivalent is count or noncount. In this sense, all nouns in modern Chinese are noncount. There are a few contexts where a noun can be quantified without a classifier, for example, in formulaic expressions handed down from classic Chinese (e.g. *jiu niu er hu zhi li* ‘the strength of nine bulls and two tigers – tremendous effort’), in parallel structures (e.g. *san tou liu bi* ‘three heads and six arms – superhuman power’), and some compact forms (*liu guo yuyan* ‘languages of six countries’), but these are exceptions rather than the norm of classifier use. The typical structure for quantifying nouns in Chinese is *numeral + classifier + noun*.

The use of classifiers in Chinese dated as early as over 3,300 years ago, as evidenced by the Oracle bone inscriptions excavated in the Yin Ruins of the Shang Dynasty (about 1300 – 1046 B.C), though technical terms such as ‘classifier’ and ‘measure word’ are only recent phenomena. Different terms have been used for classifiers over time. A popular practice was to treat them as a special class of nouns until classifiers became established as a separate word class in Chinese in the 1950s (cf. Zhou 2006), as proposed in *A Talk on Grammar in Modern Chinese* (Ding et al. 1961). The classifier was the last to have become one of the eleven word classes in Chinese because members in this category cannot function independently as sentential constituents. They typically follow a numeral, or a demonstrative pronoun such as 这 *zhe* ‘this’, 那 *na* ‘that’ or 哪 *na* ‘which’ to form numeral-classifier or demonstrative-classifier structures.

There is increasing consensus that classifiers in Chinese can be grouped into three broad categories: nominal (people, animals and objects etc), verbal (actions and events) and temporal (time), but different authors have different ideas of the subcategories for nominal and verbal classifiers. In the taxonomy based on grammatical functions and level of grammaticalization, there are three types of nominal classifiers: specialised (i.e. fully grammaticalized classifiers with no other grammatical function), concurrent (i.e. functioning either as classifiers or some other word classes), and temporary (i.e. other word classes used as classifiers on an *ad hoc* basis), while verbal classifiers are of two types: specialised (functioning only as classifiers) and borrowed (i.e. temporary borrowings from other word classes) (cf. Guo 1987). However, without access to appropriate historical corpora of Chinese, we

are not in a position to evaluate the classification based on grammaticalization. Instead, we have adopted an annotation scheme which is more appropriate for synchronic studies of classifiers in Chinese.

In the semantically-based taxonomy, there are eleven categories that have been proposed: unit classifiers, collective classifiers, standardised measure classifiers, container classifiers, species classifiers, arrangement classifiers, classifiers denoting indefinite quantities, classifiers denoting magnification, temporal classifiers, verbal classifiers, and compound classifiers (Zhu et al. 1996). In this classification, words indicating magnification (e.g. 倍 *bei* ‘times’, 成 *cheng* ‘one tenth’) are clearly numerals rather than classifiers; so are words indicating small quantities (i.e. 点/点儿 *dian/dianer* ‘a few (little)’ and 些 *xie* ‘some’), though they are sometimes treated as collective classifiers (e.g. Guo 1987: 16). Compound classifiers (e.g. 人次 *renci* ‘person-time’) are closely associated with actions and events and thus can be viewed as verbal classifiers. Hence, our Chinese corpora were annotated using a tagset that differentiates between the remaining eight semantic categories of classifiers as shown in Table 1 (cf. Xiao 2006). Note that some classifiers are polysemous because they can be used as different kinds of classifiers. For example, depending on context, 把 *ba* can be a unit classifier (e.g. *yi ba dao* ‘a knife’), an arrangement classifier (*yi ba cao* ‘a bundle of straw’), or a verbal classifier (e.g. *la ta yi ba* ‘give him a tug’). In such cases, each instance was evaluated in context to determine its category.

Classifier	POS tag	Types	Tokens
Arrangement classifier	qa	58	476
Collective classifier	ql	25	624
Container classifier	qc	38	254
Standardised measure classifier	qm	88	2,385
Species classifier	qv	8	3,181
Temporal classifier	qs	31	3,618
Unit classifier	qu	159	21,255
Verbal classifier	qt	39	2,960
Total		421	34,753

Table 1: Eight categories of classifiers in Chinese.

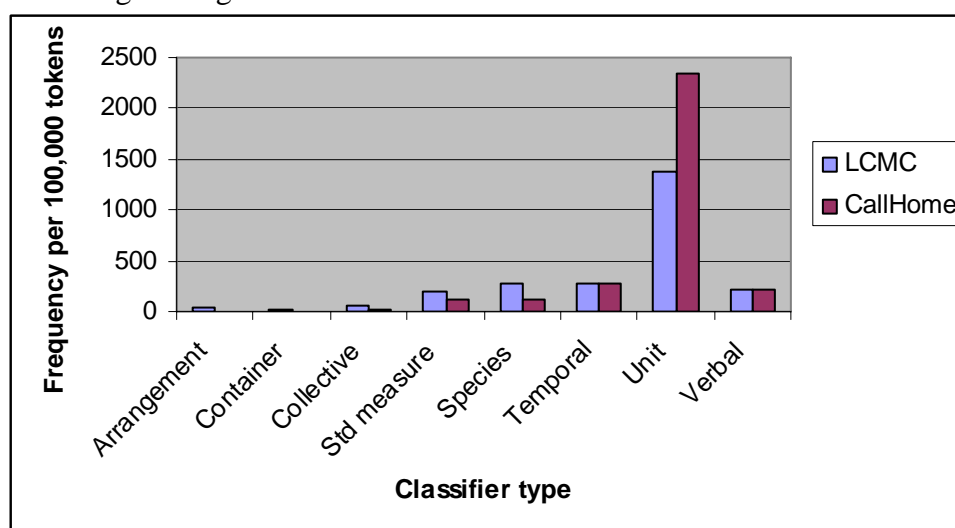


Figure 1: Normalized frequencies of different categories of classifiers in Chinese.

Figure 1 shows the normalized frequencies of classifiers of various kinds in our Chinese corpora. It is clear that of these categories, unit classifiers are by far the most common, in both writing and speech, accounting for 61.16% of total occurrences of all classifiers. Unit classifiers also contribute to over one third (37.77%) of all classifier in terms of types. The large inventory and widespread use of unit classifiers is a feature that distinguishes Chinese as a classifier language from many other languages such as English (see section 2). It is also clear in the figure that unit classifiers are considerably more common in spoken Chinese whereas species and standardised measure classifiers are more frequent in writing. Other commonly used categories include temporal and verbal classifiers, whilst container, arrangement, and collective classifiers are relatively rare (0.73%, 1.37% and 1.80% respectively).

Various genres differ in classifier use. Figure 2 shows that classifiers are most frequent in humour (R), conversation (S) and news reportage (A), with a normalized frequency of over 3,000 instances per 100,000 tokens for each of these genres. In contrast, classifiers are least common in religious writing (D), news editorial (C), academic prose (J), and news review (B), with a normalized frequency of below 2,000 instances per 100,000 tokens. Generally speaking, classifiers are more common in imaginative (K-R) writing and conversation (S) than in informative writing (A-J).

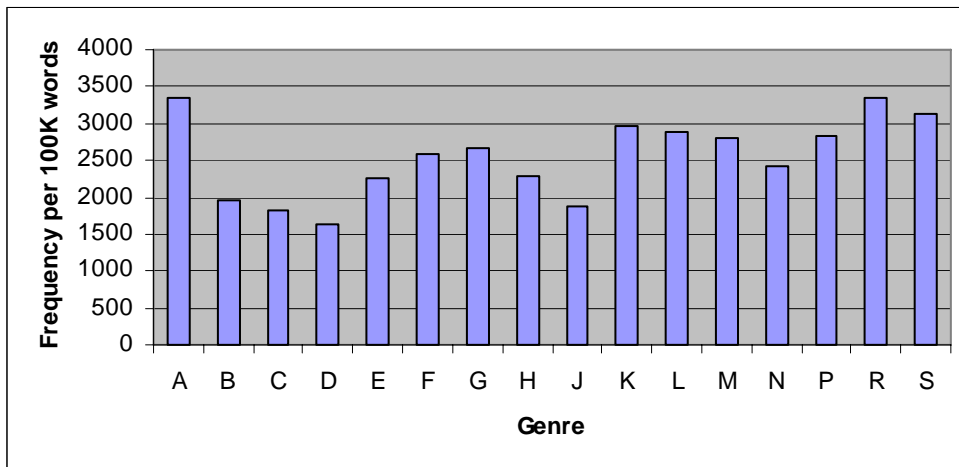


Figure 2: Distribution of Chinese classifiers across genres.

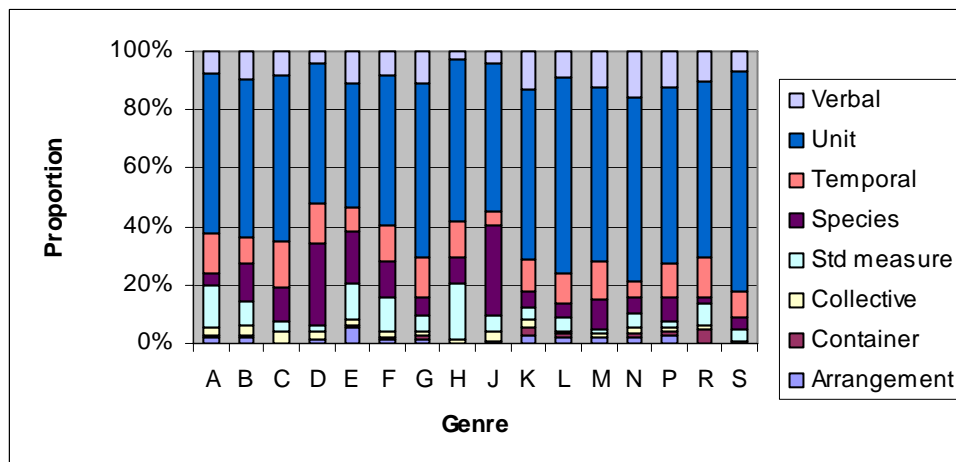


Figure 3: Proportions of different types of Chinese classifiers in each genre.

The distribution of different kinds of classifiers also varies across genres. As can be seen in Figure 3, unit classifiers are the most common category in all genres (accounting for two thirds of all classifiers) whereas container classifiers, arrangement classifiers, and collective classifiers are relatively rare in all genres. Standardised measure classifiers are most frequent in news reportage (A) and official documents (H). In addition, species classifiers are more common in informative than imaginative writing.

The most common and widely used unit classifier in Chinese is 个 *ge*, which accounts for 63.5% of unit classifiers and 38.8% of all classifier tokens. The Chinese character for the generalized classifier *ge* (i.e. 个) suggests that it is a bamboo (i.e. 竹) split into halves. As a result, 个 *ge* was initially used as a classifier to count bamboos or bamboo arrows. When a bamboo chip is used for counting, it becomes a symbol of the entity being counted. In other words, the entity loses its shape, colour, function or any other attribute and is reduced to a unit of counting. As such, *ge* can be used for any noun that does not have a specialised classifier (people or things, large or small) and can be used to replace the specialised classifiers of many nouns.¹ The widespread use of the generalized classifier *ge* is also mirrored by its collocations.² Of the 115 noun collocations, 29 refer to human beings, and 86 to non-human entities; 66 refer to concrete entities and 49 to abstract entities, including 12 related to time. The top 20 noun collocations of *ge* also show the diversity of entities which can be referred to by *ge*: *yue* ‘month’, *xingqi* ‘week’, *ren* ‘person’, *xiaoshi* ‘hour’, *dianhua* ‘phone call’, *libai* ‘week’, *zi* ‘character’, *baifendian* ‘percentage’, *difang* ‘place’, *jiaoluo* ‘corner’, *xiangmu* ‘project’, *zhongtou* ‘hour’, *wenti* ‘problem, question’, *dianfanguo* ‘rice cooker’, *nüren* ‘woman’, *zi’er* ‘character’, *lizi* ‘example’, *hezi* ‘box’, *zhaoxiangji* ‘camera’, and *dongxi* ‘stuff’. In fact, *ge* is used so frequently and extensively, especially in spoken Chinese (where it is nearly three times as frequent as in writing) and imaginative writing (nearly twice as frequent as in informative writing), that there have been observations of and arguments for the replacement of all nominal classifiers with *ge* (e.g. Li and Thompson 1980: 112; Jin and Chen 2002: 13). The complete replacement, though, is highly unlikely given that classifiers in Chinese have developed from a very small number into a complete system (cf. Sun 1996; Xue 2006), as demonstrated by the rich diversity of classifiers used in naturally occurring language.

In addition to *ge*, other commonly used unit classifiers in Chinese include, for example (in descending order of frequencies), 位 *wei* (a polite form used for persons), 条 *tiao* (for something of a long narrow shape), 张 *zhang* (for something of a thin and flat shape or with a flat surface), 名 *ming* (for persons in general), 件 *jian* (for clothes, furniture, and matters etc.), and 句 *ju* (for sentences and verbal utterances). It is clear even from these few examples that the majority of unit classifiers in Chinese do not have direct equivalents in English, because unit classifiers in Chinese have lost their lexical meanings to varying degrees, depending upon the level of grammaticalization. It is also clear that unit classifiers referring to things are closely related to shapes which are historically associated with the nouns that have given rise to these classifiers. That explains why unit classifiers and the nouns they individuate for quantification co-select each other primarily by shape. For example, *tiao* when used

as a noun in ancient Chinese meant a twig or small branch of a tree, which is of a long, narrow and flexible shape. Thus *tiao* as a classifier collocates with *jie* ‘street’, *tui* ‘leg’, *lu* ‘road’, *xian* ‘line; thread’, *he* ‘river’, and *yu* ‘fish’ etc. The use of *tiao* to quantify abstract concepts such as *guiding* ‘regulation’, *jianyi* ‘suggestion’, *faliu* ‘law’ and *xinwen* ‘news’, on the other hand, derived historically from the content word *tiao* which was related to bamboo slips for writing (cf. Jin and Chen 2002: 11). On the other hand, different classifiers can be used for the same nouns to indicate their shapes (e.g. 片 *pian* ‘flat piece’ and 颗 *ke* ‘small round piece’ for *yao* ‘medicine, pill’) or to show what the speaker thinks something looks like (e.g. 眼 *yan* ‘eye’ and 口 *kou* ‘mouth’ for *jing* ‘well’).

In relation to unit classifiers, other categories have undergone a lesser extent of desemanticization, i.e. bleaching of lexical meaning. In fact, most of them have specific lexical meanings, as can be seen in the English glosses of the examples of classifiers given below. Commonly used collective classifiers in Chinese include examples such as 套 *tao* ‘set’, 批 *pi* ‘batch’, 双 *shuang* ‘pair’, 系列 *xilie* ‘series’, 副 *fu* ‘set; pair’, and 群 *qun* ‘group, crowd’, 代 *dai* ‘generation’, 组 *zu* ‘group’, 对 *dui* ‘pair’, and 队 *dui* ‘team’. Classifiers of this category refer to entities as a group or collection. It is of interest to note that in grouping individuals together to provide a collective reference, some collective classifiers can add a layer of pragmatic meaning, especially when they refer to people. For example, 伙 *huo* ‘crowd, gang’ is more likely than 群 *qun* ‘group, crowd’ and 帮 *bang* ‘group’ to show the speaker’s negative evaluation of the group referred to.

Like unit classifiers which are largely related to the shapes of the individual items they quantify, arrangement classifiers indicate constellational arrangements. The two categories of classifiers differ in that the shapes associated with unit classifiers are natural or inherent of the objects being quantified whereas arrangement classifiers indicate the arranged or perceived shapes. For example, *yi zhang zhi* ‘a sheet of paper’ (with a unit classifier *zhang*) is different from *yi da zhi* ‘a pad of paper’ (with a collective classifier *da*) or *yi tuan zhi* ‘a ball of paper’ (with an arrangement classifier *tuan*). Frequently used arrangement classifiers include 层 *ceng* ‘tier, layer’, 堆 *dui* ‘pile’, 团 *tuan* ‘ball’, 沓 *da* ‘pad, wad’, 串 *chuan* ‘string’, 丝 *si* ‘thread’, 排 *pai* ‘row’, 把 *ba* ‘bunch, handful’, 滴 *di* ‘drop’, and 束 *shu* ‘bunch, bundle’. Both arrangement and collective classifiers provide a collective reference, but the former focus on the constellation aspect, i.e. how entities are arranged or grouped together whereas the latter do not.

Container classifiers are normally nouns denoting types of containers, which are borrowed temporarily to quantify mass or entities usually associated with them, e.g. 杯 *bei* ‘cup, glass’ for wine or drink and 碗 *wan* ‘bowl’ for food. Other common classifiers of this category include examples such as 盒 *he* ‘box, case’, 袋 *dai* ‘bag, sack’, 桶 *tong* ‘barrel, tub, bucket’, and 瓶 *ping* ‘bottle’. It is interesting to note that some nouns that denote parts of the human body or an enclosed area, e.g. 脸 *lian* ‘face’, 嘴 *zui* ‘mouth’, 头 *tou* ‘head’, 肚子 *duzi* ‘belly’, and 桌子 *zhuozi* ‘tableful’ are sometimes used as temporary classifiers. A peculiarity of classifiers of this kind is that they are more descriptive than quantifying as the preceding numeral is typically restricted to *yi* ‘one’, which is equivalent to *man* ‘full’. Consequently, it has been argued that examples like these are not to be considered as classifiers (e.g. Zhang 1994). However, as Wu (1994) and Guo (1996) observe, there are various kinds of constraints which can affect the choice of numeral for some classifiers. For example,

番 *fan* co-occurs only with *yi* ‘one’ or *ji* ‘several’, but it is a generally accepted verbal classifier.³ Furthermore, as we will see in section 2, classifiers of this kind are very similar to container classifiers ending with *-ful* in English.

Standardised measure classifiers express exact measures of various kinds in either international or local units. The most frequently used standardised measure term in our Chinese corpora is the Chinese currency unit 元 *yuan*, which is usually replaced by 块 *kuai* in spoken Chinese (28 times as frequent as in writing) and informal written genres (e.g. fiction and humour). Other commonly used measure terms include international units like 米 *mi* ‘metre’, 吨 *dun* ‘ton’, 克 *ke* ‘gram’, 美元 *meiyuan* ‘US dollar’, 厘米 *limi* ‘centimetre’, 度 *du* ‘degree’, 平方米 *pingfangmi* ‘square metre’, as well as unique measure terms in local unit, e.g. 里 *li* ‘distance of 0.5 kilometre’, 亩 *mu* ‘area of 0.165 acre’, and 斤 *jin* ‘weight of 0.5 kilogram’.

Temporal classifiers are measures of time. Some of them provide exact measures, e.g. 年 *nian* ‘year’, 天 *tian* ‘day’, 岁 *sui* ‘year of age’, 分钟 *fenzhong* ‘minute’, and 小时 *xiaoshi* ‘hour’ while others are inexact measures, e.g. 会儿 *huir* ‘a moment’, and 段 *duan* ‘a length of time’. All of them are used frequently in Chinese.

While temporal classifiers are associated with the length of time of a situation, verbal classifiers measure the count of an action or event. Chinese has a large inventory of verbal classifiers, some of which are specialised classifiers that have been fully grammaticalized while others are temporarily borrowed from nouns. Commonly used specialised verbal classifiers include 次 *ci* ‘occurrence, time(s)’, 下 *xia* ‘time(s)’, 场 *chang* ‘spell; occurrence of recreational or sports activity’, 番 *fan* ‘a course of action that takes efforts’, 下子 *xiazi* ‘time(s)’, 阵 *zhen* ‘a step of action’, 趟 *tang* ‘a return journey’, 回 *hui* ‘occasion’, 遍 *bian* ‘a course of action from the beginning to the end’, and 顿 *dun* ‘a meal; an action of criticising or reprimanding etc’. It is clear from their glosses that some specialised verbal classifiers are very widely used whereas others are restricted to certain kinds of action or event. In addition to these specialised verbal classifiers, some nouns denoting parts of the human body and instruments or media are often borrowed temporarily as verbal classifiers, e.g. 声 *sheng* ‘sound, voice’, 眼 *yan* ‘eye’, 口 *kou* ‘mouth’, 刀 *dao* ‘knife’, and 脚 *jiao* ‘foot’.

A species classifier denotes the type of entities grouped together for quantification. Common classifiers in this category include examples such as 种 *zhong* ‘kind, sort, type’, 类 *lei* ‘kind, type’, 级 *ji* ‘grade, class’, and 样 *yang* ‘kind, type’. Of these *zhong* ‘kind, sort, type’ is by far the most frequent, accounting for over 90% of all instances of species classifiers.

It was noted earlier that one of the syntactic features of classifiers in Chinese is that they cannot be used alone. They occur in numeral/demonstrative-classifier structures. Alternatively, monosyllabic classifiers can be reduplicated to function as different sentential constituents, expressing a general grammatical meaning, i.e., co-existence or repetition of entities or events, which can have different situational variants such as ‘all around’, ‘many’, ‘one by one’, and ‘continuous’ (Guo 1999).

Another syntactic feature of Chinese classifiers is that they rarely take a modifier. The modifiers they take are restricted to a very small number of adjectives indicating size and shape. For example, Lu (1987) examined 630 classifiers (including temporary borrowings), finding that 129 of them (20%) could take a premodifier, and only seven adjectives were acceptable as classifier modifiers: *da* ‘big, large’ (107), *xiao* ‘small’ (93), *man* ‘full’ (39), *zheng* ‘whole’ (34), *chang* ‘long’ (7), *hou* ‘thick’

(5), and *bao* ‘thin’ (2).⁴ Luo (1988) argued that some of the co-occurrences of classifiers with modifiers judged unacceptable in Lu (1987) could be acceptable in highly special contexts. The acceptability tests in both Lu (1987) and Luo (1988) are intuitively based. Intuitions are a powerful tool that enables one to conceive of a special context for an even most bizarre usage. In the two Chinese corpora used in this study, only 16 classifiers take a modifier, accounting for 6.2% of the total of 421. In terms of tokens, merely 0.21% of the classifiers (74 out of 34,753 tokens) are modified. Three modifiers are found in our corpora: *da* ‘big, large’, *xiao* ‘small’ and *zheng* ‘whole’, which occur 51, 16 and 7 times respectively. *Da* ‘big, large’ is not only the most frequent classifier modifier, it also co-occurs with a wide range of classifiers. We refer to modifiers like this as ‘classifier intensifiers’, which are used to reinforce the quantifier meaning of classifiers.⁵ Of the eight categories of classifiers, no instance of standardised measure term is found to take a modifier, which supports Lu’s (1987) observation that classifiers of this category cannot take a modifier. In contrast, collective classifiers, arrangement classifiers and container classifiers are most likely to take a modifier. In quantifying constructions in Chinese, a modifier normally modifies the word that follows immediately, which means that classifier modifiers are not relocated as they are in English (see section 2). Hence, *xiao* ‘small’ modifies the classifier *duan* ‘length’ in *yi xiao duan lu* ‘a short length of road; a short distance’ whereas it modifies *lu* ‘road’ in *yi duan xiao lu* ‘a length of path’. However, the contrast is not always so sharp. For example, while the adjective *da* ‘big, large’ modifies the classifier *kuai* ‘block, piece’ and the noun *shitou* ‘stone’ respectively in *yi da kuai shitou* ‘a big piece of stone’ and *yi kuai da shitou* ‘a piece of big stone’, there is not much difference in what the two phrases mean.

While the normal word order of a quantifying construction in modern Chinese is for a noun to follow a quantifying construction (i.e. numeral plus classifier), this norm can be violated to achieve special effects by focusing upon the quantity or the evaluative meaning of the noun. The inverted quantifier construction of this kind occurs typically at the end of a clause so that the end focus falls upon the quantifier. The quantifier is taken into focus in this context mainly for two reasons. One reason is that there is a contrast in enumeration of various items (*ganlanyou 20 haosheng, qu ke jidan yi zhi* ‘20 ml of olive oil, and one peeled egg’, LCMC: E). This pattern is most common in skills/trades/hobbies (E) and official documents (H). The other reason is that the quantifier is topicalized in a conjoined clause (e.g. *anpai zhili xiangmu 1789 ge, jungong xiangmu 1491 ge, jungong lü da 88.2%* ‘Of 1,689 pollution control projects which were initiated, 1,491 were completed, accounting for 88.2%’, LCMC: H). This pattern typically occurs in official documents (H) and news reportage (A). The focus on quantifiers makes these genres appear more authoritative while the archaic flavour of inverted quantifier constructions is compatible with their formal style. Inverted quantifier constructions are also found in a very different context, though much less frequently, where the numeral is restricted to *yi* ‘one’, and the noun is highly evaluative – usually deriding or negative – and becomes the focus while the quantifier functions to instantiate a member of the evaluated class (cf. Chu 2001), e.g. *huobao yi ge* ‘A funny fellow’. This pattern typically appears in speech or literary texts.

The interaction between nouns and classifiers is a complex issue. While there are cases where there is an almost one-to-one correspondence between nouns and classifiers (e.g. 门 *men* for *gongke* ‘school subject’ and 条 *tiao* for *he* ‘river’), there are also classifiers/nouns that can co-occur with nouns/classifiers very flexibly. The classifier *ge* is one such example as noted earlier, and the noun *dongxi* ‘stuff’ is

another. *Dongxi* is a noun with a rather general and vague referent. It can refer to anything, but not human beings.⁶ The vagueness in reference makes it possible to use a nominal classifier of any type for *dongxi* as attested in our corpora, for example, generalized unit classifiers such as 个 *ge*, 件 *jian*, and 份 *fen* ‘portion’, shaped-based 条 *tiao*, 张 *zhang*, and 块 *kuai*; or specialised unit classifiers for books and writings such as 本 *ben* and 篇 *pian* (then we know that *dongxi* refers to books or papers). Collective classifiers such as 套 *tao* ‘set’ can also be used; so can arrangement classifiers like 堆 *dui* ‘pile’, container classifiers like 箱子 *xiangzi* ‘box’ or 包 *bao* ‘pack’, or standardised measure terms such as 吨 *dun* ‘ton’, as well as species classifiers like 样 *yang*, 种 *zhong* and 类 *lei*. There are a number of criteria that affect the choice of classifier in Chinese.⁷ We have noted earlier that a principal criterion for the co-selection between nouns and nominal classifiers (notably unit and arrangement classifiers) is the cognitive basis of shape (Shi 2001), which is supported by collocation analysis. The second criterion is co-selection by metonymy. This means that the original lexical meanings of classifiers refer to the most salient features of the objects being classified, for example, 口 *kou* ‘mouth’ for pigs, 头 *tou* ‘head’ for cattle, 尾 *wei* ‘tail’ for fish, and 顶 *ding* ‘top’ for hats and sedan chairs etc. The third criterion is co-selection by relatedness. In this case, the original lexical meanings of classifiers refer to actions closely related to entities being classified, e.g. 包 *bao* as a verb means ‘pack’, and as a classifier it is related to the result of packing. The same can be said of classifiers such as 串 *chuan* ‘string, bunch’, 捆 *kun* ‘bundle’, and 捧 *peng* ‘double handful’. The fourth criterion is co-selection by association. This means that the original lexical meanings of classifiers refer to tools, containers, and places, etc closely associated with the entities being classified, e.g. 刀 *dao* ‘knife; cut’ as a classifier for *rou* ‘meat’, 碗 *wan* ‘bowl’ for *fan* ‘rice’, 床 *chuang* ‘bed’ for *beizi* ‘quilt’, and 幕 *mu* ‘curtain’ for *ju* ‘play, opera’. The fifth criterion is co-selection by analogue. For example, because *kuzi* ‘trousers, pants’ is counted as 条 *tiao*, this classifier is also used for *duanku* ‘shorts’ even though it is not of a long shape like a pair of trousers. Finally, co-selection by conventions is an important criterion for use of classifiers. Sometimes, the choice of a classifier has to be interpreted by following long-term linguistic conventions, because it is not always possible to track the grammaticalization path of a classifier to ascertain the relationship between its original lexical meaning with the entities being classified. For example we do not know in what way 条 *tiao* is historically related to *renming* ‘human life’; why 头 *tou* ‘head’ can be used for pigs and cattle but not rabbits or cats; why 尾 *wei* ‘tail’ can be used for fish but not peacocks or squirrels even though their tails are as salient as, if not more so than that of fish (cf. Guo 1996: 102).

We have so far given a corpus-based account of classifiers in Chinese. The section that follows will provide a parallel account of their counterparts in English before the classifiers in the two languages are compared in section 3.

2. Classifiers in English

English maintains a distinction between count and noncount nouns and distinguishes between the singular and plural forms of count nouns morphologically. Count nouns can be quantified directly by a numeral or quantifying determiner (e.g. *three books*) while noncount nouns must take a classifier when they are quantified (e.g. *a few*

pieces of bread). The typical quantifying constructions in English include a) *numeral + count noun* and b) *numeral + N1 + of + N2*.⁸ While pattern b) is well recognised as a typical quantifying structure (e.g. Allan 1977; Lehrer 1986; Brems 2003), the first noun (N1) in the structure is rarely considered as a classifier. Nouns in this slot are treated differently in English grammars. For example, they are referred to as ‘partitive nouns’ in Quirk et al. (1985: 249) and as collective nouns, unit nouns, quantifying nouns, and species nouns in Biber et al. 1999 (247–57). While the quantifying function of N1’s in such constructions has been recognised, their status has rarely been systematically questioned (cf. Brems 2003: 285). However, nouns of this kind are clearly parallels of Chinese classifiers discussed in section 2. It can be reasonably argued that they are typically classifiers in English, bearing in mind some obvious differences between the two languages which will be discussed in section 3.

Following Allan (1977), Lehrer (1986: 111) lists seven categories of classifiers in English: unit counters, fractional classifiers, number set classifiers, collective classifiers, varietal classifiers (i.e. ‘species classifiers’ in our terms), measure classifiers, and arrangement classifiers. Of these, however, the so-called fractional classifiers (e.g. *three quarters of the cake*) and number set classifiers (e.g. *many hundreds of people*) are actually quantifiers rather than classifiers, in our view. They are just like ordinary numerals (compare *many hundreds of people* and *three hundred people*). Lehrer’s (1986: 119) measure classifiers include both ‘exact and inexact measures’, which are generally recognised separately as standardised measure terms (e.g. *an acre of land*, *two pints of lager*) and quantifying nouns denoting types of container (e.g. *a cup of coffee*) (e.g. Biber et al. 1999: 252). In addition to these categories, temporal nouns such as *hour*, *day* and *week* can be used as classifiers (e.g. *200 hours of community service*) while *times* can indicate the count of an action or event (e.g. *I’ve seen it three times now*). Hence, it is clear that English has the same eight categories of classifiers as discussed in section 2 for Chinese.

The first seven categories take the form of ‘numeral (+ adjective) + N1 + of (+ adjective) + N2’ while the last is in the form of a numeral greater than two followed by *times*. However, not all instances of N1 are classifiers and it is not always possible to draw a clear-cut line between those which are classifiers and those which are the head of a noun phrase (cf. Dixon 1982: 214; Biber et al. 1999: 257). All occurrences of the pattern ‘numeral + *times*’ are not classifiers either, because such a structure can also refer to multiplied instances (e.g. *25 times greater* and *Two times one is two yeah?*). We extracted all instances of these two patterns from FLOB and BNCdemo and evaluated each concordance in context to determine whether a particular instance is a classifier and its category if it is. For the first pattern, the semantic focus was used as the primary criterion for identifying an occurrence of N1 as a classifier. If the semantic focus falls upon N2 in a noun phrase, N1 is judged as a classifier, otherwise the noun phrase has N1 as the head modified by the prepositional phrase introduced by *of* (e.g. *one word of warning*). Quantifying determiners such as *a (large) amount of*, *a couple of*, *a lot of*, *a great deal of*, and *a (small) number of* were excluded in this section. As *load(s)* and *mass(es)* behave in a way similar to quantifying determiners,⁹ they were also excluded. Each instance of the second pattern was also evaluated in context and only those indicating counts of actions or events were included.

A classifier is normally mandatory when a noncount noun is quantified while it is optional for count nouns in English.¹⁰ However, a classifier can indeed add an extra layer of meaning, for example, by indicating shape and size (e.g. *a loaf of bread* vs. *a piece of bread*) beyond its individuation or quantizing function, though some classifiers (e.g. *bit* and *piece*) have undergone a greater level of ‘desemanticization’

(Brems 2003: 290).¹¹ Consequently, it can make a difference by optionally using or omitting a classifier when a count noun is quantified (e.g. *I love a piece of orange, I I wouldn't eat a whole orange but I do like er little pieces of oranges now* FLOB: K).

While unit classifiers can also have a lexical meaning, their primary function is to individuate undifferentiated mass for quantification which cannot otherwise be quantified. Unsurprisingly, unit classifiers typically co-occur with noncount nouns. This is in spite of their co-occurrences with a small number of plural count nouns, as in *a bit of sweets, a bar of raisins, a portion of chips* and *a pair of jeans*,¹² which normally take the plural form when they express the intended meaning. In contrast to unit classifiers, collective classifiers group single entities together so as to provide 'a collective reference for separate entities' (Biber et al. 1999: 250). The items referred to collectively can be similar or different, ordered or unordered; they can be people, animals or inanimate entities. Collective classifiers typically co-occur with plural count nouns. Arrangement classifiers are similar to collective classifiers in that they also refer to a group and usually co-occur with plural count nouns. The difference between the two categories is that collective classifiers focus on the collective aspect of the items in a group whereas arrangement classifiers focus on the constellation aspect (i.e. shape) of a group, i.e. how the entities in a group are arranged (e.g. compare shapes indicated by *bunch, pile, row* and *scattering*).

Nouns denoting types of containers are often used as classifiers. A container classifier typically combines with nouns denoting what is normally, but not necessarily, contained in that kind of container, literally or metaphorically, e.g. *cup* with *coffee/tea* and *glass/bottle* with *wine/beer*. Container classifiers also include nouns ending with the suffix *-ful* (e.g. *handful, mouthful*) and *-load* for vehicle nouns (e.g. *coachload, lorry-load*). Without such suffixes these nouns are not normally construed as container classifiers (e.g. *a mouthful of food* vs. **a mouth of food*). *Handful* is perhaps the most commonly used container classifier with affixation. It is used so frequently in a metaphorical sense to refer to a small quantity that it is being grammaticalized as a quantifying determiner. The suffix *-ful* or *-load* unmistakably marks a quantifying classifier (compare *an adjacent room of stuffed creatures of the region* vs. *a roomful of grateful nuns*). As such, even typical container nouns can be affixed optionally; when an ensuing *of*-phrase is absent, the affixation is also required for the intended meaning (e.g. *he eats a big bowlful* and *I found a boxful around the back*). Container classifiers are sometimes referred to as 'irregular measures' or 'inexact measures' (cf. Lehrer 1986: 119) which, together with standardised measure terms, constitute the category of measure classifiers. In this article, however, we choose to maintain a distinction between container classifiers and standardised measure classifiers because the former are virtually limitless whereas the latter form a small closed set. Different standardised measure terms are used in measuring different kinds of things such as weight, length, area and capacity. In addition to measurement, some terms can be used as exaggerations for very small or large quantities. For example, *an ounce* is often used in negation to mean a very small amount (e.g. *not an ounce of common sense*) whereas the bare plural form *tons* without a numeral or quantifying modifier is commonly used in speech as a general exaggeration for a large quantity (e.g. *You gonna get tons of people there*) rather than as a standardised measure for weight. Container classifiers and standardised measure terms combine with noncount nouns more frequently than with plural count nouns.

Temporal classifiers measure the length of a period of time. They are different from temporal adverbials (e.g. *Simmer for 10 minutes*) in that they indicate the duration of the state of affair introduced by the *of*-phrase in the *N1 of N2* construction

(e.g. *eight hours of talks*; *two years of drought*), though both temporal classifiers and temporal adverbials can express the same idea as exemplified by the following minimal pair: *six months of preparation* vs. *prepare for months*. This distinction is justified for two reasons. On the one hand, the temporal nouns in such constructions are structurally parallel to classifiers of other categories in English (barring the verbal classifier *times*); on the other hand, they are cross-linguistically parallel to temporal classifiers in classifier languages such as Chinese (see section 2). While the sequence *numeral + times* does not take the form of *N1 of N2*, *times* is nevertheless a classifier that allows an action or event to be quantified.¹³ While *numeral + times* syntactically functions as an iterative adverbial (also referred to as aspectual adverbial), it ‘turns out to behave like a full quantified noun phrase’ (Doetjes 1997: 200) and ‘there are many indications showing that *times* can be seen as a classifier’ (Doetjes 1997: 188).

Finally, species classifiers denote types of entity referred to by the *of*-phrase. They are relevant to quantification because entities grouped together as a particular type can also be quantified. The most frequently used species classifiers in English include *sort*, *kind* and *type*.¹⁴ Classifiers such as these can co-occur with both noncount and count nouns.

It is important to note that some classifiers can belong to different categories, depending upon their meaning and entities being quantified. For example, *course* can be a collective classifier (e.g. *a course of skin treatments*) or a unit classifier (e.g. *a main course of grilled medallions of venison with garlic flavoured lentils and fried celeriac*); *line* can be a unit classifier (e.g. *one line of text*) or an arrangement classifier (e.g. *a line of parked cars*); *pack* can be a container classifier (e.g. *a pack of frozen peas*) or a collective classifier (e.g. *a pack of playing cards*); and similarly, *cup* can be a container classifier (e.g. *a cup of tea*) or a standardised measure term (e.g. *two cups of flour*).¹⁵

Classifiers of different kinds vary considerably in their productivity. Figure 4 shows the normalized frequencies of the eight categories of classifiers in FLOB and BNCdemo. As can be seen, unit classifiers are the most common category (41.7 instances per 100,000 tokens), followed by collective (32.1), container (26.8) and species classifiers (12.6). All other categories of classifiers occur less than 10 times per 100,000 tokens. Speech and writing also demonstrate noticeable differences in their use of classifiers. For example, while unit classifiers are common in both speech and writing, collective, arrangement, species and temporal classifiers are overwhelmingly more frequent in writing whereas container classifiers are much more common in speech. Classifiers of the same category also show marked contrast in productivity. For example, *bit(s)* and *piece(s)* occur 597 and 241 times respectively in FLOB and BNCdemo while more than half of the unit classifiers occur just once (e.g. *blob*, *patter*, *reel*, *stretch*, *whack*). Similarly, whilst *cup* and *bottle* are commonly used as container classifiers (356 and 108 instances respectively in the two corpora), most container classifiers are infrequent (e.g. *drum*, *lungful*, *sachet*). Table 2 gives the numbers of types and tokens of classifiers of different kinds in FLOB and BNCdemo.

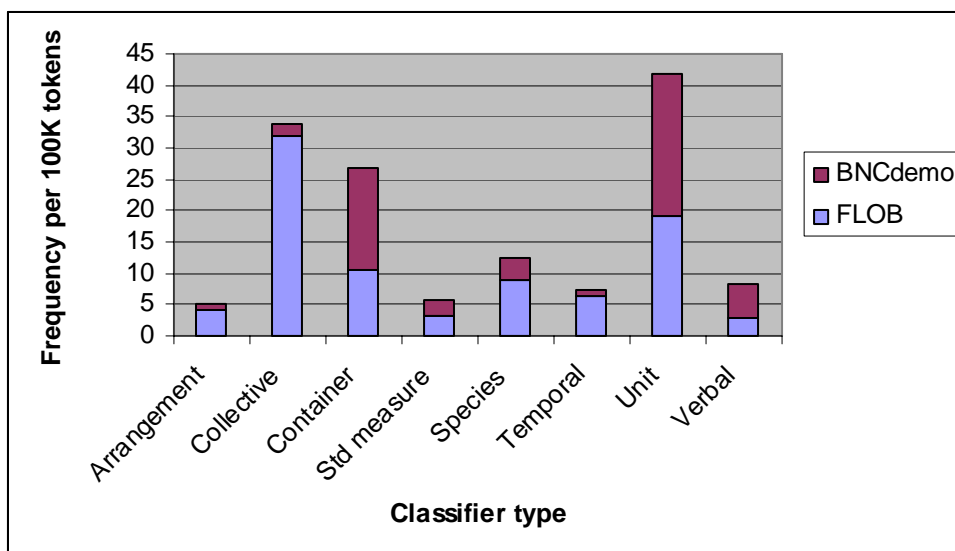


Figure 4: Normalized frequencies of classifiers in English.

Classifier	Types	Tokens
Arrangement classifier	16	108
Collective classifier	48	469
Container classifier	63	915
Standardised measure classifier	20	153
Species classifier	11	283
Temporal classifier	13	111
Unit classifier	99	1,330
Verbal classifier	1	295
Total	271	3,664

Table 2: Common classifiers in English.

Apart from productivity, classifiers of different kinds, as well as individual classifiers of the same category, also demonstrate considerable differences in collocability. Classifiers which have undergone a higher level of desemanticization are expected to show greater flexibility in collocations and to co-occur with a wider range of nouns. In terms of classifier categories, standardised measure classifiers are most limited in collocability (but see the metaphorical use of *ounce* and *tons* as noted earlier),¹⁶ followed by container classifiers, which normally combine with a noun denoting the referent entity usually associated with the type of container referred to by a container classifier (but see the highly delexicalized *handful* for an exception). As such, when we come across *a gallon of X* and *a glass of Y*, we do not have many choices for *X* and *Y*. In contrast, species classifiers such as *sort*, *kind* and *type* can combine with nouns flexibly – almost limitlessly. This is in spite of the very limited collocability of some species classifiers, for example, *brand*, *design* and *model* only collocate with nouns for manufactured goods, *breed* with animals, and *race* with human beings. Between the two extremes on a scale of collocability are unit classifiers, arrangement classifiers, and collective classifiers.

Unit classifiers in English are in a sense real classifiers because they individuate inseparable mass for quantification which cannot otherwise be quantified. Unit classifiers are also more frequent and diversified than other categories in English. However, apart from a few of them which are highly frequent and ‘characteristically general in meaning’ (Biber et al. 1999: 249), e.g. *bit* and *piece*, the majority of unit classifiers have a narrow range of nominal collocations. For example, *pinch* typically collocates with *salt* and sometimes with *sugar*, *dose* is closely related to ‘medicine’ in a literal or figurative sense, while *glimmer/glimmering* combines literally with light or metaphorically with something pleasant (e.g. *hope, interest, understanding, cheer*). In contrast, *bit* and *piece* combine with more than a hundred collocations and interact systematically with the meanings of nouns which are either concrete (*bread, paper, rock, sunshine, wood*) or abstract (e.g. *engineering, luck, research, work*).¹⁷ On the other hand, the same noun can be quantified using different unit classifiers which focus on features such as different shapes and sizes. For example, *bread* can be referred to as a *bit* (small piece), a *chunk* (large thick piece), a *hunk* (thick piece), a *loaf* (oblong shaped piece), a *lump* (small cubic piece), a *piece*, a *slab* (thick flat piece), or a *slice* (thin broad piece). Other similar unit classifiers include *bar* (thin oblong piece), *flake* (very small piece), *sheet* (thin broad piece), *slip* (narrow piece), *speck* (tiny piece), *stick* (long piece), and *strip* (long narrow piece). Each of these classifiers gives an indication of the shape and/or size of the piece in question. Modifiers of the classifier *piece* can also indicate shape (e.g. *L-shaped, triangular*) and size (e.g. *big, small*). The choice of a classifier is often a linguistic convention, which is partially motivated and partially arbitrary. Hence, while a *bit* is by definition smaller than a *piece*,¹⁸ we can have a *large bit* and a *small piece*.

Arrangement classifiers are more restricted in collocability than collective classifiers. An arrangement classifier only combines with nouns whose reference entities can be arranged, physically or conceptually, into the shape denoted by the classifier (e.g. *300 bales of straw, six ropes of pearls*). However, some arrangement classifiers, when they are used to indicate figurative configurations, can combine widely with nouns, e.g. *a bunch of flowers (keys, girls, blokes), a bundle of (feathers, handouts, laughs, nerves), a heap of shit (clothes, trouble, worries), a pile of books (leaves, debts, queries, words), and a string of beads (lovers, objections, convictions)*. Collective classifiers also differ from each other in collocability. For example, *flight* as a collective classifier can only be used to quantify stairs/steps or something that is flying (e.g. *hawks, swallows, wide ducks and moths*). In contrast, some collective classifiers such as *group, range, series* are very flexible in collocations. While the majority of collective classifiers are used neutrally, some can express an affective meaning. For example, of the collective classifiers that can refer to people (e.g. *army, band, bevy, brood, clutch, company, crowd, file, flock, gang, group, host, mob, pack, party, posse, stream, swarm, team*), collocations show that *gang, mob, pack* and *swarm* clearly show a negative evaluation.

Brems (2003: 285) observes that classifiers ‘are typically associated with informal registers.’ But this observation is only partly supported by our data. Figure 5 shows the distribution of classifiers across 16 genres covered in our corpora. As can be seen, classifiers are generally more common in imaginative writing, which is normally considered to be more informal than informative writing (A-J). However, it is quite surprising to find that conversation (S), which is supposed to be more informal than other individual genres, shows the lowest frequency (53.9 instances per 100,000 tokens) of all genres covered in our English corpora, much lower than the average frequency for written genres (87.4 instances per 100,000 tokens). A closer

look at concordances revealed that not all kinds of classifiers are associated with informal genres. As can be seen in Figure 6, which shows the proportions of different categories of classifiers in each genre, while container classifiers and unit classifiers are generally more frequent in imaginative writing and conversation, collective classifiers and standard measure classifiers are much more common in informative writing (A-J). Temporal classifiers are common in genres such as news texts (A-C), popular reading (F), humour (R) and fiction (K-P) while other kinds of classifiers do not appear to show a clear distribution pattern because of their low overall frequencies.

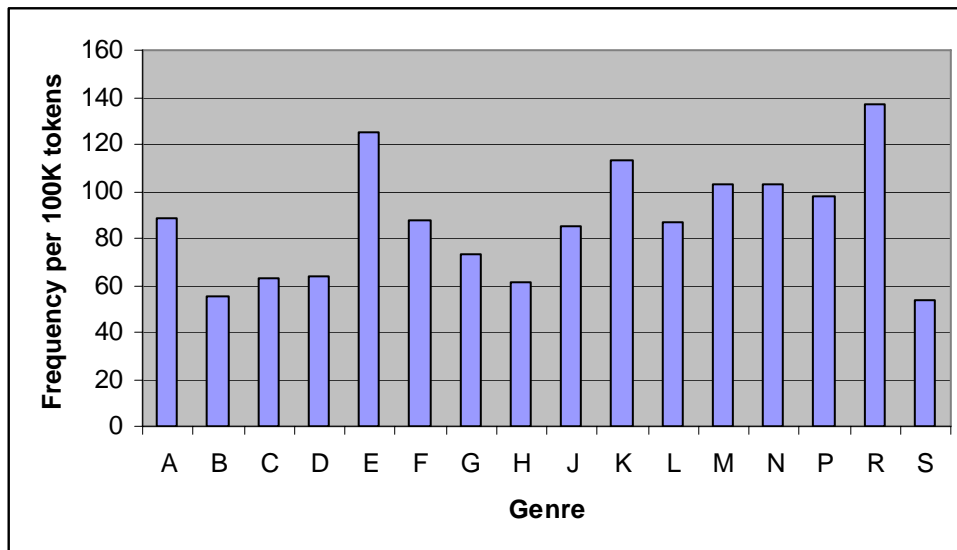


Figure 5: Distribution of classifiers across genres in English.

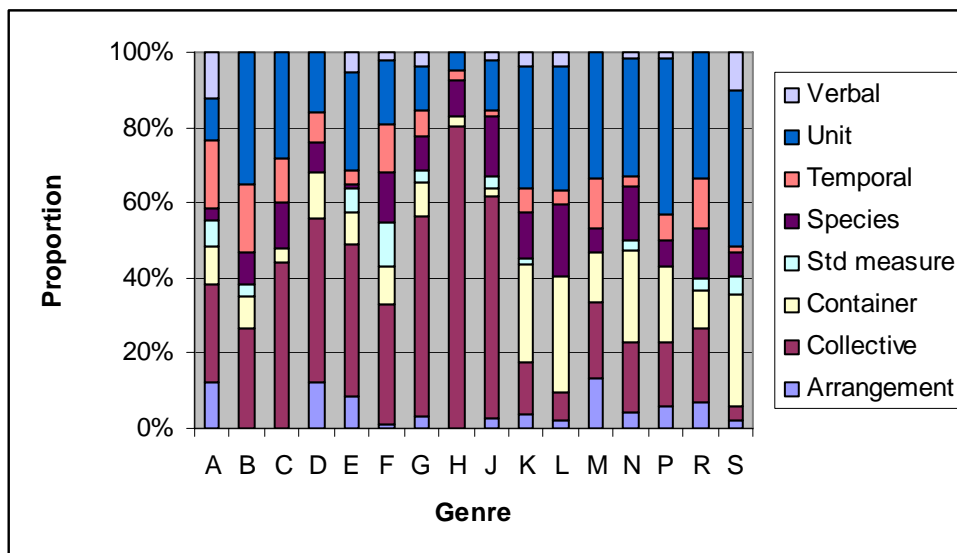


Figure 6: Proportions of different types of classifiers in English.

With the exceptions of the verbal classifier *times* and standardised measure terms, which are not preceded by a modifier, classifiers of other categories can take a premodifier. As in Chinese, the typical and most common premodifiers are classifier

intensifiers (e.g. *full, great, huge, large, little, small, whole*). Evaluative qualifiers of N2 are frequently relocated to precede classifiers (e.g. *a late-night cup of coffee, an old piece of machinery*). Swear words can also be relocated (e.g. *a lump of bloody lard* vs. *a bloody lump of lard; a fucking pint of beer* vs. *a bit of fucking chocolate*). However, when a modifier of N2 does not have an evaluative meaning, it cannot be relocated (e.g. *a glass of red wine; a handful of German scholars*); neither can a modifier be relocated from N2 to N1 if N1 already has a pre-modifier (e.g. *a large mug of strong tea, a strong cup of tea* vs. **a large strong cup of tea*). Similarly, a modifier preceding a classifier is not necessarily relocated from N2 (e.g. *an earthen pitcher of clean water*). However, modifier relocation is largely pragmatically motivated and cannot be predicted systematically (cf. Brems 2003: 300).

3. A Contrastive Analysis of Classifiers in English and Chinese

The discussions in the two sections above suggest that classifiers in English, while they may not be termed as such in grammar books, are clearly parallels of classifiers in Chinese in spite of the differences which will be discussed shortly. In contrastive research, therefore, one should not be confused by the different terms used for the same phenomenon in the languages under consideration. This section considers classifiers in English and Chinese in a contrastive context.

Before classifiers were established as a separate word class in Chinese in the 1950s, they were treated as a special group of nouns just as they are in English today. Indeed, classifiers and nouns are so closely interwoven that no firm line can be drawn between the two out of context. On the one hand, as noted in section 2, a large number of commonly used classifiers in Chinese nowadays are temporary borrowings from nouns, which function as classifiers on an *ad hoc* basis; on the other hand, some classifiers can be used in a way similar to nouns. Even in context, it is not always possible to make a clear distinction between classifiers and nouns (see Li 2000). In spite of the interwoven relationships, nevertheless, classifiers were separated from nouns to become a word class of their own in Chinese because of their mandatory grammatical status, whereas their counterparts in English are conventionally not considered as a separate word class because they are only required for noncount nouns. In other words, they are optional for count nouns even though they can affect meaning. In contrast, Chinese is a language that does not make a morphological distinction between single and plural nouns, or between count and noncount nouns at all, which means that an appropriate classifier is required for all nouns in Chinese, with a few exceptions as noted in section 1. It is clear that classifiers have a much wider scope of use in Chinese than in English. Consequently, it is hardly surprising that classifiers are 29 times as common in Chinese as in English in their overall frequency of use (2,251 and 88 instances per 100,000 words respectively).

English and Chinese have the same eight categories of classifiers, but as can be seen in Figure 7, which shows their proportions in terms of tokens, unit classifiers are predominant in Chinese whereas container classifiers and collective classifiers are significantly more common in English. While unit classifiers are also the most frequent category in English, their normalized frequency (42 instances per 100,000 tokens) is much lower than that in Chinese (1,866 instances per 100,000 tokens). In terms of types, Chinese has a greater number of unit classifiers, standardised measure classifiers, arrangement classifiers and verbal classifiers whereas English uses more collective classifiers and container classifiers (Figure 8).

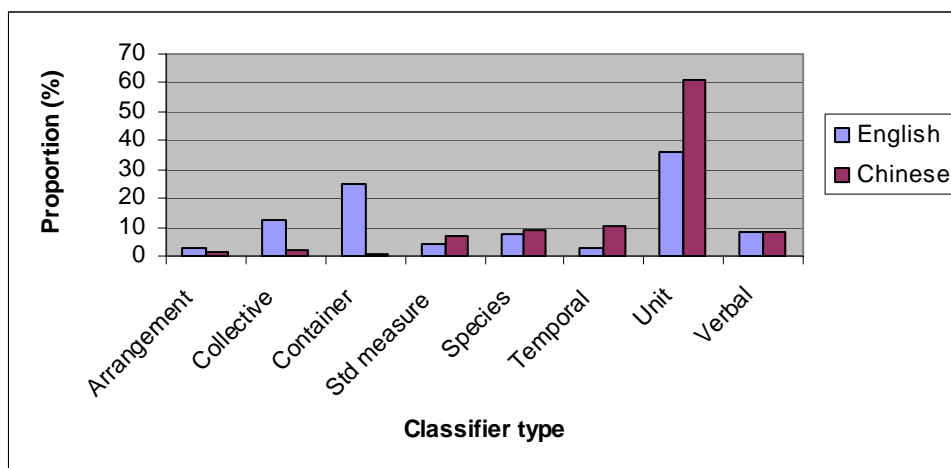


Figure 7: Proportions of eight categories of classifiers in terms of tokens.

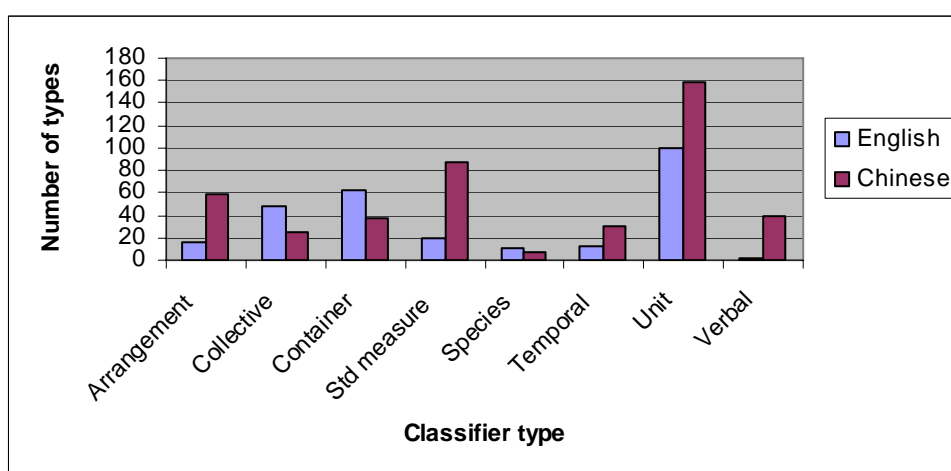


Figure 8: Numbers of eight categories of classifiers in terms of types.

In spite of the quantitative differences in the types and tokens of classifiers used, English and Chinese do not differ much qualitatively in their use of classifiers. Of the eight categories of classifiers, the most noticeable difference lies in unit classifiers, because their individuation is mandatory for all nouns in Chinese whereas they are only required for noncount nouns in English. Other categories of classifiers are qualitatively more similar than different in the two languages. They have full lexical meanings and can find their counterparts in the other language, though they are likely to be known by different terms. An interesting coincidence is that a special group of nouns temporarily borrowed as container classifiers in Chinese, which cannot take numerals other than *yi* 'one' (with the meaning of *man* 'full'), are parallel to container classifiers ending with the suffix *-ful* in English (e.g. *handful*, *fistful*, *armful*, *mouthful*, *eyeful*, *earful*, *lungful*), except that such English classifiers are not restricted to singular forms (e.g. *armfuls of wet blankets and clothes*, FLOB: N).

On the other hand, there is an important difference in the way actions and events are quantified in the two languages. In Chinese, there are some fully fledged verbal classifiers and a large number of *ad hoc* verbal classifiers which are borrowed from nouns denoting parts of the human body and instruments or media, whereas in

English, the verbal classifier *times* and adverbs *once* and *twice* are used to indicate the count of actions or events. In addition, English relies heavily upon light verb constructions, which consist of a light verb and a verbal action noun (e.g. *have a look, give the car a push, fired two shots*), to approximate the quantifying function of temporarily borrowed verbal classifiers in Chinese. While there are some variations in frequencies of use of standard measure terms, species classifiers and temporal classifiers in English and Chinese, these categories do not differ much in their behaviours in the two languages.

A common feature of arrangement classifiers and unit classifiers in Chinese and English is that they are largely motivated by the cognitive basis of shape, as exemplified by unit classifiers 块 *kuai*, 条 *tiao* and 张 *zhang* in Chinese and arrangement classifiers *bunch, pile* and *row* in English. Some classifiers are also motivated pragmatically. For example, English classifiers like *gang, mob* and *pack* usually refer to a group of people the speaker does not approve of, which differentiates them from more neutral collective classifiers such as *crowd* and *group*. Similarly, in Chinese, some collective classifiers (e.g. 伙 *huo*) are habitually negative in evaluation; and so are some verbal classifiers (e.g. 通 *tong*), whereas unit classifiers such as 位 *wei* can only be used for respectable people.

There are a number of syntactic differences in the use of classifiers in English and Chinese. Firstly, English classifiers as a special group of nouns have singular and plural forms while their counterparts in Chinese do not. Secondly, the majority of monosyllabic classifiers in Chinese can be reduplicated whereas classifiers in English cannot. Thirdly, the numeral *yi* 'one' in quantifying constructions can be omitted in Chinese if they function as objects (e.g. *xie (yi-)feng xin* 'write a letter' and *wo hai you (yi-)ge wenti* 'I have another question'), but quantifying determiners and numerals in English cannot. Fourthly, inverted quantifying constructions are found in Chinese but not in English. Finally, while they do not regularly take a modifier in both languages, classifiers take a considerably greater variety of modifiers in English than in Chinese. Classifier modifiers in Chinese are largely classifier intensifiers, which emphasize the large or small quantity or amount. There are two major types of classifier modifiers in English, i.e. classifier intensifiers like their counterparts in Chinese, and evaluative qualifiers relocated from the nouns being quantified. No such relocation occurs with classifier modifiers in Chinese.

Finally, classifiers show similar distribution patterns across genres in English and Chinese (see Figure 9), with noticeable exceptions in conversations (S) and news reportage (A), which have greater proportions of classifiers in Chinese, and in skills/trades/hobbies (E) and humour (R), which have higher proportions of classifiers in English. Classifiers are considerably more common in conversations and news reportage in Chinese than in English largely because of the overwhelming proportion of unit classifiers in Chinese, and in the case of news reportage, also because of the frequent use of standardised measure terms. Classifiers are much more common in skills/trades/hobbies (E) and humour (R) in English simply because these are two genres where classifiers are most frequent. With regard to classifier types, container classifiers are very common in conversations and imaginative writing in English but are rare in all genres in Chinese.

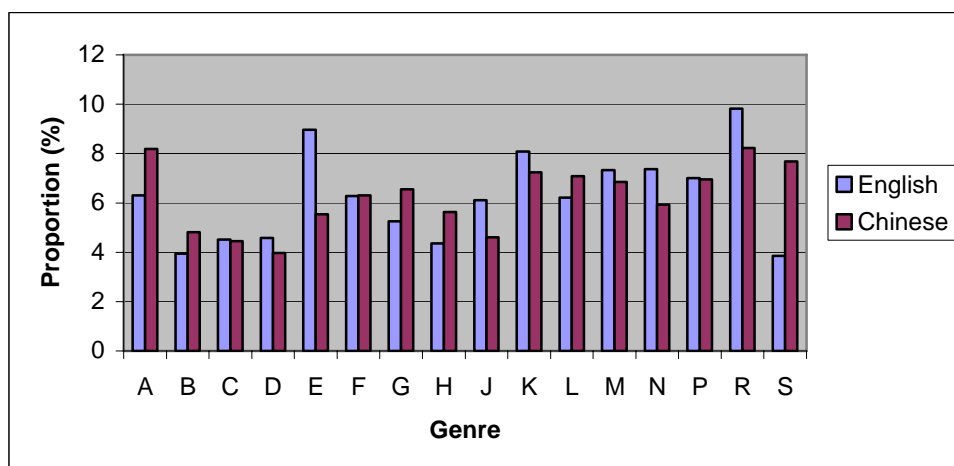


Figure 9: Proportions of classifiers across genres.

4. Interaction of Quantifying Constructions with Progressives

We have so far explored quantifying constructions in Chinese and English, focusing upon classifiers. This section examines an important aspect of quantifiers, which interact closely with aspect as they are syntactically associated with arguments and adjuncts. On the one hand, quantified internal arguments always occur in telic situations while on the other hand some viewpoint aspects tend to shun quantifiers (see Xiao and McEnery 2004a). A quantifier, be it a numeral, a quantifying determiner, or a numeral-classifier structure discussed in sections 2 and 3, has the aspectual function of delimiting a situation when it interacts with telic verbs (Xiao and McEnery 2004b). A situation with a final end point is typically presented from a temporal perspective that views the situation as a whole, using a perfective viewpoint aspect such as the perfect or the simple past. The progressive as an imperfective viewpoint focuses on a subinterval while ignoring the boundaries and is therefore in clash with quantifiers. Semantically, a quantifying construction usually conveys new information and becomes the semantic focus whereas the progressive focuses upon the ongoing nature of an event.¹⁹ Consequently, the two are reasonably expected to shun each other (Hatav 1989; Abe 1998; Filip 2005).²⁰ However, our corpus data shows that the progressive does co-occur (albeit only occasionally) with quantifying constructions, in both English and Chinese. For example, quantified progressives account for 0.64% of total progressive forms in FLOB and 0.47% in BNCdemo; they take up 1.9% of total occurrences of quantified object nouns in FLOB and 3% in BNCdemo. The interaction between the progressive and quantifiers has so far rarely been explored. In this section, we examine the conditions under which the progressive interacts with quantifier constructions, specifically quantified objects and adverbials indicating a temporal duration (see Xiao and McEnery 2006 for further discussion of durational adverbials).

Let us first consider quantified objects in progressive forms. A total of 18 instances of quantified objects in progressive forms are found FLOB and 116 instances in BNCdemo. Five of these appear in future progressives (two from FLOB and three from BNCdemo; e.g. *So daddy, you'll be consuming four hot cross buns!* BNC: KBW) and ten occur in perfect progressives (all from BNCdemo; e.g. *Somebody's been eating a lot of those haven't they?* BNC: KBH). A future

progressive focuses on its futurate meaning while a perfect progressive form has a perfect focus. Neither of these constructions purely indicates progressiveness. Hence, it is quite natural for them to take a quantified object.²¹ A progressive form in English (i.e. *BE + -ing*) does not always signal the progressive or ongoing nature of a situation as its name suggests. Rather, it has a number of other usages including functions such as indicating an anticipated future happening, or a habitual or iterative situation, and as a politeness strategy (Leech 1987). Of the 119 remaining instances in our English data, 17 express a futurate meaning (one from FLOB and 16 from BNCdemo; e.g. *I'm taking two young ladies out on Sunday for lunch*, BNC: KDM) while 10 express a habitual meaning (two from FLOB and eight from BNCdemo; e.g. *at the end of the decade, Skae was giving two lectures weekly*, FLOB: H).²²

An analysis of the remaining 92 instances of progressive forms marking progressiveness reveals that the quantified object is most commonly (58.7% of the time) a noncount noun (e.g. *Flanked by four huge minions, he was making a lot of noise*, FLOB: R; *So they're making ninety five quid?* BNC: KC2). As a noncount noun refers to a mass (time and money, etc) that is indivisible, it can be said that the objects in these cases are internally quantified and are treated as a whole. Another interesting feature of progressive forms taking a quantified object is that the verb in the progressive form is frequently (27.2% of the time) a stative verb (e.g. *MPs are having a spot of telephone trouble*, FLOB: B; *Do you want to know why I'm wearing two pairs of knickers because*, BNC: KCE). As only a dynamic event can have subintervals that can be singled out as a focus to indicate the progressiveness of the event, the stative progressive in these cases signal temporariness rather than progressiveness of a stative situation (cf. Leech 1987). It can be said that the quantified objects are 'not linked to the Incremental Theme' in stative progressives (Filip 2004).²³

There is, however, a small proportion (14.1%) of dynamic progressives in our English corpora that take a quantified count noun object, e.g. *POLICE are investigating a number of offences relating to cars which occurred over the weekend* (FLOB: A), and *The fact is now I mean where like two years ago she was taking two and three exams at a time* (BNC: KBF). In spite of their low frequency in attested data, examples such as these are perfectly acceptable, because the progressive as an imperfective aspect forces the addressee to focus on the internal structure of an event rather than its result (cf. Delmonte 1997), which would otherwise become the focus if the event was presented with a perfective aspect. An event designated by a quantified object allows for either a simultaneous or sequential reading, a distinction which is irrelevant in perfective aspects. But when a quantified object interacts with the progressive, its sequential reading is ruled out (cf. Filip 2004). This means that in our examples cited above, the offences were being investigated simultaneously and the exams were being taken at one time within the time frame under observation. As with the internal quantification of noncount nouns noted earlier, the sub-events denoted by a quantified object are viewed as a single 'meta-event' (Abe 1998), or as a 'collective event' (Ogihara 1990) in the progressive, otherwise the use of progressive forms would be out of place (cf. Yeom 2003: 148).

Now let us consider the interaction of the progressive with durational adverbials. According to Mourelatos (1981) and Hatav (1989), the progressive does not co-occur with durational adverbials because the two have conflicting focuses. End points are specified by durationals but they are excluded in progressives. The most comprehensive account of this interaction is probably provided in Mittwoch (1988: 224–33), according to which 'the progressive in its primary sense (the "imperfective"

one) is incompatible in simple past tense sentences with durationals like *for 2 hours* that give an exact measurement of time' (Mittwoch 1988: 224). This means that progressive forms with a futurate reading can interact freely with durational adverbials. It is not clear why Mittwoch specified the simple past tense as a condition. Probably because she thought that past progressive forms describe what happened in the past and cannot express a futurate meaning. But in fact they can. They express a relative future meaning, as illustrated in her example *John was working for 2 hours* (Mittwoch 1988: 224). Neither is it clear why a distinction is maintained between durationals 'that give an exact measurement of time' and 'vague and usually hyperbolic durationals (e.g. *for hours, for ages*) now that both vague and exact durationals (e.g. *for half an hour, for 5 hours*) are found in non-futurate progressives. It is also unconvincing, in our view, to attribute the interaction of non-futurate progressives with durational adverbials (vague or exact) to 'a survival from an earlier stage of language (Mittwoch 1988: 232).

In FLOB only seven instances of progressive forms were found that interact with a *for*-adverbial. According to Xiao and McEnery (2006), not all *for*-adverbials indicate temporal duration. They may have a scope reading, especially when *for*-adverbials appear in the clause-initial position or interact with negation, superlative adjectives, ordinal numbers, and words such as *only*. Two of the occurrences in FLOB are of this kind (i.e. *Burgundy was having its hottest summer for years*; FLOB: F; *For many years the denomination was drifting from its roots, and only recently can we speak of an evangelical resurgence*, FLOB: D). All of the other five instances have a futurate reading, as Mittwoch (1988) would have us expect, e.g. *The Empress of Britain "is calling at Los Angeles for five minutes, and then calling at New York for five minutes", he told reporters* (FLOB: G). Note that there are some special cases of futurate progressives where the speech moment is contained in the time frame denoted by an durational adverbial. For example, *I'm – sort of – babysitting for an hour or so* (FLOB: P) was uttered by someone in the middle of *an hour or so* rather than before the babysitting event started. Mittwoch (1988: 227) uses the term 'predictive or predeterminate sense' to refer to this special futurate sense.

However, the results from BNCdemo are both as expected and unexpected. Of the 48 co-occurrences of progressive forms with durational adverbials in BNCdemo, 37 have a futurate or predeterminate reading, accounting for 77.8% of the total, while the remaining 11 instances simply defy such a reading in their context. In that latter group of examples, durational adverbials give either a vague (e.g. *I says I was getting that for years Berty*, BNC: KPJ) or exact (e.g. *He was doing so well for three months*, BNC: KC9) measurement of time; the subject can be either human (*you said she was having vitamins for a while didn't you?* BNC: KDM) or inanimate (e.g. *I've done it since I was, well it's getting on for fifteen years now*, BNC: KC0);²⁴ and such co-occurrences were found in sentences of both past (e.g. *Margaret was living with that lad for ten years weren't she?* BNC: KCG) and present tense (e.g. *This new car it's going on now for seven weeks*, BNC: KP1). Mittwoch (1988: 232) notes that non-futurate progressive sentences taking durational adverbials 'which focus on the shortness of the time taken are much less acceptable.' However, our data shows that durational adverbials focusing on shortness of time are equally acceptable, e.g. *I'm just looking after her for a while* (FLOB: P) and (talking about a twelve-pound big baby) *Yeah, she's only coming up for six weeks, mm, a monster* (BNC: KBH). Temporal duration is a mental or psychological concept (Xiao and McEnery 2004b), hence *a while* and *six weeks* can be considered as long or short durations, depending on one's psychological expectation. In the above two examples, the 'attitude

diminishers' (Quirk et al. 1985: 598) *just* and *only* clearly indicate that the speakers think that *a while* and *six weeks* are short durations.

Such usages account for over 20% of the total instances. They are found in speech of different people with both southern and northern accents of British English, and from a range of occupations including housewife and shop assistant as well as teacher, engineer, administrator, consultant, and advertising executive. As such, they cannot be simply dismissed as slips of tongue. Neither can they be simplistically attributed to a revival of earlier usages as Mittwoch (1988) suggested. A common feature of these examples is that they all occur in casual conversations where errors such as *I says* and *weren't she* are also found. They are non-standard usages.

The above discussion shows that the progressive forms interacting with quantified objects and durational adverbials usually express a futurate or predeterminate meaning. It also appears that English is more tolerant of non-futurate progressives taking a quantified object than those taking a durational adverbial. The latter only occurs in non-standard English while the former typically co-occurs with a noncount noun or a stative verb, or expresses the simultaneous reading of a collective event.

In comparison with English, the progressive marked by *zai* (or *zhengzai*) in Chinese only marks a dynamic situation in progress. It cannot express a futurate or habitual meaning. Nor can a stative verb be used in Chinese progressives. Then can the progressive *zai* (*zhengzai*) interact with quantifying constructions in Chinese?

Liu (1994) observes that the Chinese progressive is incompatible with quantifying adverbials that indicate a period of time (e.g. **ta zhengzai chi yi xiaoshi de pingguo* 'He is eating apples for an hour') or frequency (e.g. **ta zhengzai chi liang ci pingguo* 'He is eating apples twice'), but the interaction between the progressive and quantified direct objects is more complicated. According to Liu (1994), a quantified object is acceptable in progressives only when the quantifier expresses a quantity of one and the noun is referential (e.g. *ta zhengzai zhao yi ben shu* 'He is looking for a book'), because in this case the quantifier is referential (i.e. with a certain identifiable referent) rather than quantificational. Quantifiers indicating a quantity of more than one are unacceptable (e.g. **ta zhengzai zhao san ben shu* 'He is looking for three books') because they are purely quantificational, while attributive nouns are unacceptable because they cancel the referential reading of quantifiers which indicate a quantity of one (e.g. **ta zhengzai chi yi ge pingguo* 'He is eating an apple'). However, attributive nouns taking an adjectival modifier can have a referential reading, and are thus acceptable in progressives when the quantifier is one (e.g. *ta zhengzai chi yi ge you-da-you-hong de pingguo* 'He is eating a big red apple').

Liu's (1994) account of the interaction of Chinese progressives with quantifiers is interesting, and it can account for the majority of the phenomena in the attested language data represented in our corpora as well. Indeed, as Liu would have us expect, frequency adverbials are not found to co-occur with Chinese progressives. Temporal adverbials are not monolithic. They can indicate duration, period, and range, etc. (Xiao and McEnery 2006), but only those with a range meaning can interact with progressives (e.g. *zhe liang tian hai zai dao shicha ne* '(I'm) still overcoming the jet lag these two days', CallHome). While we cannot say that what is absent in a corpus is not possible in language, its absence at least suggests that it is untypical or marginalized at best if not unacceptable at all. Hence, Liu's (1994) observations of the interaction of Chinese progressives with quantifying adverbials are largely supported by our data. Nevertheless, her assertions about quantified objects are too categorical. While they are rare, quantified objects do occur in progressives, even in cases where

the quantified object is an attributive noun with no modifier (e.g. *dangshi Xiao Meng ye zai kan yi duan xinwen* ‘Xiao Meng was also reading a piece of news at that moment’, LCMC: G), or where the quantity denoted by the quantifier is greater than one (e.g. *wo ziji zai zuomo yixie shiqing la* ‘I’m thinking over a couple of things myself’, CallHome). The interaction of the progressive with quantified objects is not confined to casual speech; it is found in formal writing as well (e.g. *zhifa jigou zhengzai shencha 50 ren* ‘The law enforcement agency is examining 50 people’, *People’s Daily* 20/01/2000).

Our corpus data indicates that only a very small proportion of progressives marked by *zai* and *zhengzai* take a quantified object (2.1% and 0.7% in LCMC and CallHome respectively). The low frequency of quantified objects in progressives is as expected. On the one hand, the Chinese progressive can only focus upon the ongoing nature of an event, which is in clash with the semantic focus of quantifiers. On the other hand, as quantifiers function to provide a final end point to a situation, they tend to co-occur with perfectives but shun imperfectives. Of the co-occurrences of quantified objects with progressives found in our corpora, the most common type (roughly 69%) is *yi + classifier* ‘one’, especially when it is followed by a modified object noun. As Liu (1994) observes, the Chinese quantifier *yi* can have a referential (i.e. ‘a certain’) rather than quantificational reading (i.e. ‘one’), which is similar to English indefinite articles. That explains why *yi* is the most common numeral in quantified objects co-occurring with progressives. Because of the light weight of the referential *yi*, however, the sequence of *yi + classifier* is usually dropped when the quantified noun does not take an attributive modifier.

In relation to progressive forms in English, Chinese progressives marked by *zai* or *zhengzai* appear to be even less tolerable, if at all, to durational adverbials – at least no such co-occurrence is found in our corpora – though they can take temporal adverbials with a range meaning. Adverbials of this latter kind differ from durational adverbials in that they provide a time frame within which a situation is presented rather than indicate the duration of the situation. Progressives also interact with quantified objects less frequently in Chinese than in English for two reasons. First, unlike indefinite articles in English, the article-like *yi* is not mandatory in Chinese. Second, while noncount nouns and stative verbs account for 85% of co-occurrences of English progressives and quantified objects, the two do not apply in Chinese. On the one hand, Chinese does not make a distinction between count and noncount nouns – in a sense all nouns in Chinese are noncount because a classifier is required when they are quantified, while on the other hand, stative verbs do not occur in Chinese progressives at all. But as in English, the interaction of Chinese progressives with quantified objects only allows a simultaneous reading, i.e. the multiple events denoted by a quantified object are considered as a collective event.

5. Conclusions

In this article we have explored quantifying constructions in English and Chinese on the basis of comparable corpora that cover a range of genres, focusing on classifiers and the interaction between progressives and quantifying constructions in the two languages. In answer to our first research question, it is clear that Chinese employs numeral-classifier constructions obligatorily in quantification whereas in English a classifier is only required when noncount nouns are quantified; count nouns in English can be quantified directly by a numeral or quantifying determiner. This cross-

linguistic difference exists simply because Chinese is a non-inflectional language whereas nouns in English inflect for plurality morphologically.

As for the second research question, all of the eight semantic categories of classifiers exist in both Chinese and English. Classifiers in the two languages differ in a number of ways. For example, classifiers are significantly more common in Chinese; unit classifiers and verbal classifiers are characteristic of Chinese whereas collective classifiers are more diversified in English; classifiers in English and Chinese display a some language-specific syntactic differences; there are also some slight differences in the distribution of various categories of classifiers across genres in the two language. Nevertheless, these differences are largely quantitative rather than qualitative. Classifiers are motivated cognitively, pragmatically, and conventionally in both English and Chinese. In other words, classifiers in the two languages are less different from each other than their different terms in current use would suggest.

Regarding the third question, quantifiers can indeed interact with progressive forms in English, but such interactions are typically restricted to non-canonical futurate use of progressive forms, stative progressives, simultaneous reading of collective events, internally quantified objects, and a non-standard style. In contrast, because of their narrower scope of use, Chinese progressives are grammatically less tolerable to quantifiers than their English counterparts. Progressives in Chinese can interact with quantified objects, but this interaction is marginalized, with the quantifier typically restricted to *yi* 'one'. The other type of quantifying construction under consideration, durational adverbials, are extremely rare, if acceptable at all, in Chinese progressives.

Acknowledgements

We are grateful to the UK Economic and Social Research Council (ESRC) for supporting our project "Contrasting English and Chinese" (RES-000-23-0553), on which the work presented in this article is undertaken.

References

- Abe, Y. (1998) Quantifier scope and aspect interpretation. *Researching and Verifying An Advanced Theory of Human Language 2*, 365–88.
- Allan, K. (1977) Classifiers. *Language 53*, 281–311.
- Biber, D. (1988) *Variation Across Speech and Writing* (Cambridge: Cambridge University Press).
- Brems, L. (2003) Measure noun constructions: An instance of semantically-driven grammaticalization. *International Journal of Corpus Linguistics 8/2*, 283–312.
- Biber, D., Johansson S., Leech G., Conrad S. and Finegan, E. (1999) *Longman Grammar of Spoken and Written English* (London: Longman).
- Chu, Z. (2001) 'Ming + shuliang' yuxu yu zhuyi jiaodian (The word order 'noun + numeral-classifier' and the focus of attention). *Chinese Language 2001/5*, 411–17.
- Delmonte, R. (1997) Lexical representations, event structure and quantification. *Quaderni Patavini di Linguistica 15*, 39–93.
- Ding, S., Lü, S., Li, R. et al. (1961) *Yufa Jianghua (Talks on Grammar)* (Beijing: Commercial Press).

- Dixon, R. (1982) *Where Have All the Adjectives Gone? and Other Essays in Semantics and Syntax* (Berlin: Mouton).
- Doetjes, J. (1997) *Quantifiers and selection. On the distribution of quantifying expressions in French, Dutch and English*. PhD dissertation, Leiden University.
- Filip, H. (2005) The telicity parameter revisited. *Semantics and Linguistic Theory XIV* (Ithaca: CLC Publications).
- Guo, J. (1999) Zai tan liangci chongdie de yufa yiyi (Reanalysis of grammatical meaning of classifier reduplications). *Chinese Language Learning* 1999/4, 4–9
- Guo, X. (1987) *Xiandai Hanyu Liangci Shouce (A Handbook of Classifiers in Modern Chinese)* (Beijing: Peace Press of China).
- Guo, X. (1996) Tantan wuliangci dui qian dapei shuci de yuyi xuanze (On the semantic choice of collocation of nominal classifier and numeral). *Journal of Remin University of China* 1996/3, 98–103.
- Hatav, G. (1989) Aspects, Aktionsarten, and the time line. *Linguistics* 27, 487–516.
- Jin, F. and Chen, G. (2002) Hanyu liangci de yufahua (Grammaticalization of classifiers in Chinese). *Journal of Tsinghua University (Philosophy and Social Sciences Edition)* 2002/1, 8–14
- Hundt, M., Sand, A. and Siemund, R. (1998) *Manual of Information to Accompany the Freiburg-LOB Corpus of British English* (Freiburg: University of Freiburg).
- Leech, G. (1987) *Meaning and the English Verb* [2nd ed.] (London: Longman).
- Lehrer, A. (1986) English classifier constructions. *Lingua* 68, 109–148.
- Lepore, E. and Ludwig, K. (2003) Outline for a truth-conditional semantics for tense. In Q. Smith and A. Jokic (eds.) *Tense, Time and Reference* (Cambridge, MA: MIT Press), 49–105.
- Li, N. and Thompson, A. (1980) *Mandarin Chinese: A Functional Reference Grammar* (Berkeley: University of California Press).
- Li, Y. (2000) Liangci yu shuci, mingci de niujie (Fuzzy boundaries of classifiers with numerals and nouns). *Language Teaching and Linguistic Studies* 2000/3, 50–58.
- Liu, X. (1994) Hanyu shuliangci de yuyi fenbian ji jingxingshi dongcizu zhong shuliangci de shiyong (Semantic differentiation of classifiers in Chinese and use of classifiers in progressive verbal phrases). *Chinese Teaching in the World* 1994/4, 10–17.
- Lu, J. (1987) Shuliangci zhongjian charu xingrongci qingkuang kaocha (A survey of numeral-classifiers with an inserted adjective). *Language Teaching and Linguistic Studies* 1987/4, 333–42.
- Luo, Y. (1988) Guanyu shuliangci zhongjian charu xingrongci qingkuang de buchong kaocha (A supplemental survey of numeral-classifiers with an inserted adjective). *Chinese Language Learning* 1988/4, 7–12.
- McEnery, A., Xiao, Z. and Mo, L. (2003) Aspect marking in English and Chinese: Using the Lancaster Corpus of Mandarin Chinese for contrastive language study. *Literary and Linguistic Computing* 18/4, 361–78.
- Mittwoch, A. (1988) Aspects of English aspect: On the interaction of perfect, progressive and durational phrases. *Linguistics and Philosophy* 11, 203–254.
- Mourelatos, A. (1981) Events, processes, and states. In P. Tedeschi and A. Zaenen (eds.) *Tense and Aspect* (New York: Academic Press), 191–212.
- Ogihara, T. (1990) The semantics of the progressive and the perfect in English. In H. Kamp (ed.) *Tense and Aspect in English*. DYANA Deliverable R2.3.A.

- Quirk, R., Svartvik, J., Leech, G. and Greenbaum, S. (1985) *A Comprehensive grammar of the English language* (London: Longman).
- Shi, Y. (2001) Biao wuti xingzhuang de liangci de renzhi jichu (The cognitive basis of classifiers indicating shapes of objects). *Language Teaching and Linguistic Studies* 2001/1, 34–41.
- Sun, R. (1996) Guanyu liangci ‘ge-hua’ lun de sikao (Reflections on the overuse of ‘ge’). *Journal of Yunnan Normal University (Philosophy and Social Science Ed.)* 1994/1, 70–74.
- Wu, Y. (1994) Shuliang you xianzhi de shuliang jiegou (Quantifying constructions with restricted numerals). *Language Teaching and Linguistic Studies* 1999/4, 52–68.
- Xiao, Z. (2006) Using corpora to study classifiers in Mandarin Chinese. Talk given at the Working Groups Meeting of COST Action A31: Stability and Adaptation of Classification Systems in a Cross-Cultural Perspective. 8–10 December 2006, Humboldt-Universität, Berlin.
- Xiao, Z. and McEnery, A. (2004a) *Aspect in Mandarin Chinese: A corpus-based study* (Amsterdam: John Benjamins).
- Xiao, Z. and McEnery, A. (2004b) A corpus-based two-level model of situation aspect. *Journal of Linguistics* 40/2, 325–63.
- Xiao, Z. and McEnery, A. (2006) Can completive and durative adverbials function as tests for telicity? Evidence from English and Chinese. *Corpus Linguistics and Linguistic Theory* 2/1, 1–21.
- Xing, F. (1993) Xiandai Hanyu shuliangci xitong zhong de ‘ban’ and ‘shuang’ (‘Half’ and ‘double’ in the numeral-classifier system of modern Chinese). *Language Teaching and Linguistic Studies* 1993/4.
- Xue, J. (2006) Liangci ‘ge-hua’ wenti guanjian (Problems concerning the overuse of ‘ge’). *Chinese Language Learning* 2006/5, 22–27.
- Yeom, J. (2003) The semantics of the English progressive and the imperfective paradox. *Language and Information* 7/2, 139–161.
- Zhang, K. (1994) Guanyu hanyu liangci de jixian wenti (On the issue of demarcation line for classifiers in Chinese). *Lexicographical Studies* 1994/3, 103–107.
- Zhou, C. (2006) 20-shiji 90-niandai yilai de xiandai Hanyu liangci yanjiu zongshu (A review of research in classifiers in modern Chinese since the 1990s). *Journal of Hubei TV University* 2006/3, 43–45.
- Zhu, X. and Wang, H. (1994) Xiandai Hanyu liangci yu mingci de zilei huafen (Subcategorization of classifiers and nouns in modern Chinese). In *Papers in Computational Linguistics* (Beijing: Institute of Computational Linguistics, Peking University).

Notes

¹ For example, the specialised classifier for *fangjian* ‘room’ is 间 *jian*, but it is also possible to use *ge*. The replacement of a specialised classifier with the generalized classifier *ge*, if felicitous at all, generally makes the style more colloquial.

² Collocations here refer to the co-occurring words within a window span of five words on both left and right of the node words, with a z-score above 3.0 and a minimal co-occurring frequency of 5 in our corpora.

³ The idiomatic expression *sanfan-wuci* ‘again and again, repeatedly’ is an exception.

⁴ The figure in the brackets indicates the number of classifiers that were judged as acceptable in taking the specified modifier.

⁵ Adjectives indicating shape (e.g. *chang* ‘long’, *hou* ‘thick’, *bao* ‘thin’), when they are used as classifier modifiers, also intensify the quantifier meaning of classifiers.

⁶ It is an insult to the addressee to say that someone is a *dongxi*, or is not a *dongxi*.

⁷ Only standard Mandarin Chinese is considered in this article. There are variations arising from dialects, diachronic change and creative use of classifiers in literary works.

⁸ The term *numeral* here should be understood broadly as including cardinal numbers and indefinite articles *a/an*. For easy manipulation of data, quantifying determiners such as *a great deal* and *a lot of* are excluded from our analysis.

⁹ Brems (2003: 309) demonstrates that *loads of* is very similar to regular quantifiers like *a lot of/lots of* in the extent of grammaticalization.

¹⁰ Quirk et al. (1985: 248) observe that a classifier can be omitted when a noncount noun is ‘reclassified’ as a count noun, as in *Two coffees, please* (see also Lehrer 1986: 110). It is of interest to note that the reclassified use of *tea*, *coffee* and *beer* typically refers to a portion in the usual container (i.e. *two coffees* means *two cups of coffee*), but *wine* is rarely used in this way, the plural of which usually refers to different kinds or brands instead of portions.

¹¹ Brems (2003) observes that *bunch of* and *heaps of* are under a process of being grammaticalized as a quantifier.

¹² Lehrer (1986: 118) views *pair* as a collective classifier. However, *pair* is considered here as a unit classifier because it is typically used to quantify *pluralia tantum* nouns (e.g. *scissors* and *pants*) which actually refer to one object. Cross-linguistically, a unit classifier is used for their equivalents in Chinese (e.g. 把 *ba* for *jiandao* ‘scissors’ and 条 *tiao* for *kuzi* ‘pants’).

¹³ Apart from the verbal classifier *times*, adverbs *once* and *twice*, as well as light verb constructions such as *take a look* and *give a push*, can also quantify actions and events.

¹⁴ Instances of *sort of* and *kind of* used as hedges are excluded.

¹⁵ As a standardised measure term, one cup equals to 8 oz. in the U.S. or 10 oz. in the British Commonwealth countries (cf. Lehrer 1986: 120).

¹⁶ The verbal classifier *times* does not take the form of the prototypical classifier structure in English and is irrelevant to this discussion.

¹⁷ Lehrer (1986: 115f) argues that ‘*Interest* comes in items, not pieces. *An item of interest*, **a piece of interest*.’ But in our view, neither *piece* nor *item* in her examples is a classifier. An item of interest means an item that is interesting (e.g. *The background section of the report has a number of items of interest on company taxation, employment law and the market for building construction*, FLOB: E). *Piece* can indeed combine with *interest*, but as a head noun instead of a classifier (e.g. *We know she is a member of the Philatelic Writers’ Society, and we hope that Quill might reprint her message as a piece of general interest to philatelic writers and editors*, FLOB: E). Our data shows that the unit classifier for *interest* is *wealth*, *flick* or *glimmer*.

¹⁸ *Bit* can refer to a small piece or a small quantity, so we can say *a bit of fresh air* but not *a piece of fresh air*.

¹⁹ It is understandable that the progressive can co-occur naturally with a quantified noun denoting a single object, as in *He was reading a paperback* (FLOB: N). Hence, quantifiers denoting a single entity (e.g. *a*, *an* and *one*) are not counted as quantifiers in this section.

²⁰ Mittwoch (1988: 226) argues that ‘certain sentences can only occur in the perfective mode, specifically sentences containing activity verbs with measure phrase objects or with durational phrases (which are also measure phrases) or with ordinary objects quantified by a numeral greater than one.’

²¹ Like the perfect, the perfect progressive is fully compatible with durational adverbials probably because ‘PROG is not in the immediate scope of the durational’ – the progressive and perfect components ‘are fused together so tightly that not even a negative can separate them’ (Mittwoch 1988: 240).

²² Following Lepore and Ludwig (2003: 74–76), we view sentences with frequency adverbials (e.g. *he was drinking two pints of whisky every twenty-four hours*, FLOB: G) as habitual sentences.

²³ According to Mittwoch (1988: 229), there is ‘no truth-conditional difference in meaning’ between the progressive and non-progressive forms of semantically stative verbs (e.g. *average, carry, have, hold* and *wear*) – the two forms are ‘in free variation’.

²⁴ Mittwoch (1988: 224) uses the example *It was raining for 2 hours* to rule out the futurate or predeterminate reading because raining is typically not predetermined.