

## **People, Things and Stuff: General Nouns in Spoken Mandarin\***

Yung-O Biq

*National Taiwan Normal University*

This is a corpus-based study of general nouns, *ren* 'people', *shi(qing)* 'matter', and *dongxi* 'thing, object', as they are used in spoken Mandarin in contemporary Taiwan. Our examination focuses on the co-occurrences and collocations of these general nouns with other linguistic elements. Our investigation shows that the three lexical items display different tendencies in designating referential specificity. They also manifest different extents to which each can form fixed expressions and stabilized constructions with other linguistic elements. Finally, the behavior of these general nouns also demonstrates how language use in interaction motivates and reinforces the conventionalization of meaning construction.

Key words: general nouns, corpus-based lexico-grammar, collocation, phrasal expression, construction

### **1. Introduction**

The relationship between discourse, grammar, and lexicon is a focus of attention in recent linguistic studies. On the one hand, the emergent view of grammar maintains that grammar is but the sediment of the repeated linguistic choices motivated by recurrent communicative needs that speakers make in social interaction (Hopper, 1988, 1998). Frequency, grammaticalization, and the impact of interactive discourse on grammar are thus closely examined notions in this approach (e.g., Bybee & Hopper, 2001; Ochs, Schegloff, and Thompson, 1996). On the other hand, in the corpus-based approach, since Sinclair's observation of the idiom principle equilibrating the open choice principle (Sinclair, 1991; Erman & Warren, 2000), the phraseological and constructional aspects of how words are strung together have been intensively investigated (e.g., Biber et al., 1999; Cowie, 1998; Hunston & Francis, 2000; Moon, 1998; Partington, 1998; Stubbs, 2001).

Following the recent interest in Chinese linguistics in probing lexico-grammar as it is manifested in various types of corpus (e.g., Biq, 2004, to appear; 方 [Fang], 2002; C. Huang, 2003; S. Huang, 1999, 2000; Liu, 2001; 陶 [Tao] 2000, 2003a, Tao 2003b; 張&方 [Zhang & Fang], 1996; 鄒 [Zou], 2001), this paper proposes a study of how general nouns behave in spoken Mandarin. Integrating the theoretical and methodological contentions of the emergentist approach and the corpus-based

---

\* The author would like to acknowledge the support by a grant from the National Science Council of Taiwan (NSC 92-2411-H-003-037) for the research reported in this paper.

approach, we hope to gain more understanding of how semantically and functionally similar lexical items are differentiated in terms of the ways they are combined with other linguistic elements in actual use. In other words, this study aims to explore the syntagmatic aspect of the meaning construction of paradigmatically comparable linguistic elements. In the following, Section 2 reviews literature on the subject and identifies the lexical items to be studied. Section 3 presents the analysis and the discussion. Section 4 is the conclusion.

## 2. Background

### 2.1 Literature review

General nouns refer to ‘a small set of nouns having generalized reference within the major noun classes’ (Halliday & Hasan, 1976: 274). Examples from English include *people*, *person*, *man*, *woman* (for human nouns), *creature* (for non-human animate nouns), *thing* (for inanimate concrete count nouns), *stuff* (for inanimate concrete mass nouns), *matter* (for inanimate abstract nouns), *move* (for action nouns), *place* (for place nouns), and *question* and *idea* (for fact nouns). At first sight, general nouns are by definition lexical items that contain the least amount of substantial meaning. However, as words are never used in isolation but always appear in association with other words, general nouns are worth our attention because of the company they have and the context in which they appear. In fact, it is the general noun phrase rather than the general noun that interests us (Partington, 1998). For example, general nouns, when combined with a specific determiner, such as *the thing*, are very similar to a reference item (e.g., *The thing doesn’t work*) (Halliday & Hasan, 1976). In addition to its referring function, a general noun (phrase) can also serve as the complement of a reference item (e.g., *He’s the man*). In fact, in the Hallidayan framework, general nouns constitute a special class of vocabulary for cohesion, because a general noun is itself a borderline case between a lexical item (member of an open set) and a grammatical item (member of a closed system) (Halliday & Hasan, 1976: 274). While both referring and characterizing are ideational functions, general nouns (phrases) may also have textual functions. They are often used in texts to lend their “noun-ness” as a support for other words’ (Sinclair, 1999:172). Thompson (1988) calls this type of use “empty noun” since the “other” words, such as the attributive adjective, function predicatively (e.g., *It’s the same thing*). Mahlberg (2003) identifies three aspects of the support function of general nouns in texts: giving emphasis, adding information in passing and providing an introduction. Furthermore, some expressions consisting of general nouns, such as *this point*, are used as a retrospective label in text that packages a stretch of prior

discourse, presents it as given fact, and directs how the upcoming text should be oriented with respect to this given information (Francis, 1994). A general noun (phrase) may also have interpersonal meanings. Francis (1994), for example, argues that *this move* ‘encodes the writer’s acceptance’ that whatever the political decision that is being referred to in the text by *this move* ‘is definitely on the agenda’ (p. 88). In many cases, however, it is the modifier in a general noun phrase that indicates the writer’s attitude when the noun merely serves as a retrospective label for textual cohesiveness (e.g., *these insatiable people*).

While general nouns often display collocational relationship with other elements in a noun phrase – determiner, attributive adjectives, etc., they have also been found to co-occur with elements outside NP, forming phrasal expressions of various lengths and structural complexities. Lenk (2000), for example, identifies the ways the high-frequency English noun, *time*, is used in “stabilized expressions” that are not specifically associated with any pragmatic or organizational function, e.g., *all this/that/the time*, *give/spend/waste time*. On the other hand, expressions that are associated with pragmatic or organizational functions abound. For example, in spoken English *thing* occurs frequently in the phrasal expression, *the thing is*, which is often used to begin an utterance or a turn (Biber et al. 1999).

## 2.2 Target items and database

As lexical items that are more likely to occur in spontaneous types of communication, general nouns (phrases) have a more prominent profile in unplanned discourse than in planned discourse. In this paper, we will investigate how Chinese general nouns are used in conversation. Our spoken corpus consists of about 10 hours of audio-recordings of naturally occurring casual conversation and radio talk show programs in Taiwanese Mandarin. We will examine the “companies” that each general noun keeps, both within and outside of the noun phrase. We will investigate to see if these general nouns have developed stabilized collocates or have even formed “prefabricated” phrases/constructions or idiomatic expressions with other linguistic elements.

We need to first identify the target items for our investigation. While the concept of general nouns is easy to establish, whether a specific lexical item is a general noun is often a difficult decision. In Chinese, for example, it is not clear whether *wenti* ‘question, problem, point’ and *fangmian* ‘aspect, point’ are general nouns or not: they are frequently used to refer to what has been talked about in prior discourse, just like *dian* ‘point.’ However, it seems that as lexical items they each carry “more” meaning than *dian*. The question of whether a lexical item is a general

noun or not is also related to the text type one investigates. *Fangmian*, for example, is a frequently used device for textual organization in written and/or formal Chinese but does not occur with comparable frequency in Mandarin conversation. A contrastive study of how general nouns (phrases) are used in written Chinese and spoken Mandarin would tell us about the genre factors that influence the speaker/writer's lexical choices and eventually the construction of the meaning of a lexical item.

For the practical reason of limiting the scope of the study to reasonable bounds while given that this study is probably the first of its kind in probing Chinese general nouns from the corpus-lexico-grammatical perspective, we take the conservative approach and investigate only three lexical items, *ren* 'person, people,' *dongxi* 'thing,' and *shi(qing)* 'matter,' which should all be non-controversial as prime instances of general noun in Chinese.<sup>1</sup> Another reason for why these three (instead of some other) items are chosen is that all of them have a (relatively) large number of token occurrences in our spoken Mandarin database.

### 3. Analysis and discussion

#### 3.1 Analytical hypothesis and procedures

The analytical procedure for this study is as follows. We first analyze the internal structure of the general noun phrase, since we believe the elements that can co-occur with the head general noun can tell us most about what this general noun is like. We also examine other collocating relationships beyond NP and check if there are any stabilized expressions and/or constructions.

Given that a general noun does not have much substantial meaning to specify the properties of the referent, it can be hypothesized that when it is used to refer, the general noun phrase is used either when the referent is intended to remain vague/underspecified, or when the identification of the referent is expressed by the modifier portion in the general noun phrase and the general head noun serves only as a "carrier" for that identification. Of course, it is a matter of degree from vague/underspecified to specific identification. All the colligational elements (i.e., grammatical elements that can co-occur with a noun in an NP) are resources for nominal identification: the measure word, and various modifying expressions including possessive pronouns, attributive adjectives, restrictive adjectives, and relative clauses, etc.<sup>2</sup>

---

<sup>1</sup> Naturally, words that contain these lexical items as morpheme but express more specific meanings, such as *nanren* 'man', *bingren* 'patient', *renjia* 'other people', *rensheng* 'life', are excluded from consideration.

<sup>2</sup> Although an important part of NP, determiners (e.g., *zhe* and *na*) will not be considered in our analysis

In the case of vague/underspecified identification, one can expect occurrences of bare NPs, i.e., NP consisting of only the general noun itself. Moreover, the general noun used with a general measure word also serves the purpose. Sortal measure words such as *zhong* ‘kind’ and marker of indefinite quantity, *xie*, convey vagueness in the identification of the referent. In the case of more specific identification, one can expect the use of specific (non-sortal) measure words, various kinds of modifying expressions, and a combination of measure words and modifying expressions.

In our case analysis, the data will be divided into several NP-internal types: (1) bare N, (2) MW (measure word) + N (with or without other modifying expressions), and (3) modifier + N (with or without MW). The general nouns are found to be frequently preceded by the question word, *shenme* ‘what,’ in the data. As a versatile grammatical word, *shenme* can be used either as a question word suggesting an interrogative, as a quantifier marking totality in certain constructions, or as a hedge suggesting indeterminate reference (cf. Biq, 1990). In none of these functions the word provides further specification about the referent of the general noun. Thus, for our purpose here, we group the combination, *shenme* + N, into the first type, “bare N.” As for the relationship between general nouns and measure words, we will check which individual measure words are preferred collocates and whether they are of the general or specific type. We also check the extent to which these general nouns take modifying expressions.

Beyond NP-internal relations, we further examine, at the clause level, the grammatical roles (A, S, or O) that these general nouns take and if any, the verbs that they are associated with. Finally, at the discourse level, we check if there are any stabilized expressions and any pragmatic functions associated with them.

In the following, we discuss *dongxi*, *shi(qing)* before *ren*, since *dongxi* and *shi(qing)* are more similar to each other than to *ren* in terms of their behavior.

### 3.2 DONGXI

There are 220 tokens of *dongxi* in our data. One third of them collocate with a MW (with or without other modifying expressions). The measure words that dominate the collocation relation are *xie* (n= 29, 40% of the 72MW tokens), *ge* (n=26, 36%), and *zhong* (n=15, 21%). Overall, general MWs (n= 44, 61%) (i.e., *xie* and *zhong*) are preferred to the specific MWs (n= 28, 39%) (i.e., *ge* and other minor ones).

---

since we believe the information determiners provide is mainly concerned with the discourse role the nominal referent plays rather than how and why the general noun is used to represent this referent.

Consider the following examples:<sup>3</sup>

- (1) 而且我媽會給我一些水果或者一些東西帶上來，  
 Erqie wo ma hui gei wo yi xie shuiguo huozhe yi xie dongxi dai shang lai  
 Further my mother will give me one PL fruit or one PL DONGXI bring up  
 come  
 ‘Besides, my mom usually gives me some fruits or some (other) stuff to  
 take along when I come back (up to Taipei).’
- (2) 裡面好像最有特色的就是這個東西，  
 Limian haoxiang zui you tese de jiu shi zhe ge dongxi  
 Inside seem most have characteristic NOM just be this GE DONGXI  
 ‘It seems that the most distinctive one in there is this one.’
- (3) 小朋友本來就喜歡這種東西啊。  
 Xiaopengyou benlai jiu xihuan zhe zhong dongxi a  
 Children to-begin-with just like this ZHONG DONGXI PRT  
 ‘All kids like this kind of stuff.’

There are 84 tokens (38% of total) of *dongxi* that collocate with some form of a modifier but without any MW. However, with or without MW, almost half of the *dongxi* tokens (n= 104, 47%) take modification. The large number of the modifying type is significant in that the modifier provides further, specific information about the referred object. Consider the following examples of modification:

- (4) 有很多東西泰國保存得比我們好。  
 You hen duo dongxi Taiguo baocun de bi women hao  
 Have very many DONGXI Thailand preserve COMP COM we good  
 ‘There are a lot of things that Thailand preserves better than we do.’
- (5) 我不是在吃飯就是在弄班上的東西，不然就在寫字條。  
 Wo bu shi zai chifan jiu shi zai nong ban shang de dongxi, buran jiu zai xie  
 zitiao  
 I NEG be PROG eat then be PROG do class on NOM DONGXI otherwise  
 then PROG write note  
 ‘I am either eating or handling the affairs in our class, or writing notes.’

<sup>3</sup> These examples from the conversation data have been edited for ease of reading. Thus, features typical of spontaneous speech, such as repetition, restart, or pause, have been eliminated. Capitalized in the gloss are key words (i.e., REN, DONGXI, SHI, SHIQING) and measure words (i.e., GE, ZHONG, etc.). The following is the list of other gloss conventions: 3S – third person singular, COM – comparative, COMP – verbal complement, COMPL – completive, DUR –durative, EXP – experiential, PROG – progressive, NEG – negation, NOM – nominalizer, PRT – sentence final particle, Q – question marker.

- (6) 我們台灣的不是都是流行日本人家已經過期的東西嗎?

Women Taiwan de bu shi dou shi liuxing Riben renjia yijing guoqi de dongxi ma

We Taiwan NOM NEG be all be in-fashion Japan other-people already out-of-fashion NOM DONGXI

‘Don’t we Taiwanese always fantasize about the outdated fad in Japan?’

Over half of the *dongxi* tokens serve as the object of a verb. There are, however, no particular verbs that are associated with *dongxi*.

As expected, we find that *dongxi* and elements co-occurring with it often express referential underspecificity. The underspecificity can either arise from the context, i.e., the referent is intended to be vague in its identity, or can be attributed to the speaker’s choice of bypassing the more specific noun in characterizing the referent. Consider the following examples. In (7), the referent of *dongxi* is intended to remain underspecified given the context.

- (7) 現在大概不能拿重的東西<sup>a</sup>。

Xianzai dagai buneng na zhong de dongxi a

Now probably NEG can take heavy NOM DONGXI PRT

‘Now, it is probably not okay (for him) to carry heavy stuff.’

However, in the following example, one can argue that other, more specific, nouns such as *yinliao* ‘drink’ or *kafei* ‘coffee,’ could have been used after the verb, *he* ‘drink.’

- (8) (about a fancy coffee shop in Tokyo)

而且它連坐在那邊喝東西的時間都要算錢。

*Erqie ta lian zuo zai naban he dongxi de shijian dou yao suan qian*

Further it even sit at there drink DONGXI NOM time all want count money

‘Moreover, it even charges money for the time people spend sitting there drinking something.’

The following example illustrates the same point. A more specific (and probably more appropriate) noun after *zuo* ‘cook’, in this context, could have been *cai* ‘food’ or *fan* ‘meal.’

- (9) 我本人不挑剔的結果我也不太會做東西。

Wo benren bu tiaoti de jieguo wo ye bu tai hui zuo dongxi

I myself NEG picky NOM result I also NEG too can do DONGXI  
 ‘While I am not a picky person (about food), I don’t cook very well either.’

Furthermore, referential underspecificity is also expressed in elements co-occurring with *dongxi*. Indeed, the measure word for indefinite quantity, *xie*, is a good example. When *xie* is combined with *dongxi*, the speaker is being vague about both the quantity and properties of the referred object. Consider the following example:

(10) 而且我媽會給我一些水果或者一些東西帶上來， (same as (1))

While in the MW tabulation *ge* is counted as indexing specificity with respect to the referent of *dongxi*, it has to be admitted that in some examples the specificity is rather elusive. The *ge* in the ‘V-ge-N’ construction, for example, does not contribute specificity to the N but rather to the activity denoted by the verb (Biq, to appear). In the following example, the MW *ge* co-occurring with *dongxi* does not express the meaning ‘one thing.’ Rather, *chi ge dongxi*, as a ‘V-ge-N’ construction, expresses something close to “eat a small quantity quickly”, i.e., ‘take a quick bite.’

(11) 補完習，然後去吃個東西。然後就可以打包，去唱歌。

Bu wan xi, ranhou qu chi ge dongxi. Ranhou jiu keyi dabao, qu change  
 Cram-class-over then go eat GE DONGXI afterward then may take-out  
 go sing

‘After cram school, we can grab something to eat, pack up the leftover  
 and then go (to the KTV) to sing.’

Other vagueness expressions that co-occur with *dongxi* within the NP include modifying expressions such as *zhilei(de)* ‘of the kind of’ and *you(yi)dian* ‘somehow, somewhat.’ Consider the following examples:

(12) 而且自己也蠻少買那種袋袋之類的東西。

Erqie ziji ye man shao mai na zhong daidai zhilei de dongxi  
 further (my)self also quite little buy that ZHONG bag kind NOM  
 DONGXI

‘Besides, I myself rarely buy things like that kind of bags, either.’

(13) 然後這整條街都是賣那種，有點古早味，日本風味的東西。

Ranhou zhe zheng tiao jie dou shi mai na zhong, youdian guzaowei, riben  
 fengwei de dongxi

Then this whole strip street all be sell that ZHONG somehow  
old-fashioned Japan style NOM DONGXI  
'And the whole street is selling stuff that are sort of old-fashioned and  
Japanese-style.'

What these vagueness expressions suggest is that the speaker does not fully commit to what s/he is saying (quantity in the *xie* case, categorization in the *zhilei* case, properties in the *you(yi)dian* case), which is just the same as when s/he uses *dongxi* to name the referent, s/he is suggesting that s/he is not committing him-/herself to what the thing exactly is except that it is a "thing!"

In some extreme cases, where the modifying portion consists of nothing but a noun, the modified head noun, *dongxi*, seems most vacuous in meaning. Consider the following examples:

- (14) 他說他，一開始就會教文法，他就是一開始就會講文法的東西，  
Ta shuo ta, yi kaishi jiu hui jiao wenfa, ta jiu shi yi kaishi jiu hui jiang  
wenfa de dongxi  
3S say 3S once start then will teach grammar 3S just be once start then  
will talk grammar NOM DONGXI  
'He said that he will start with grammar, that he will talk about  
grammatical things right at the beginning.'
- (15) 因為他要考試所以他就讀很多理論的東西，  
yinwei ta yao kaoshi suoyi ta jiu du hen duo lilun de dongxi  
because 3S want exam so 3S then read very many theory NOM DONGXI  
'Since he is going to take the examination, he is reading a lot of  
theoretical stuff.'
- (16) 他們男生，就不是很去表達那種情緒的東西，  
Tamen nansheng, jiu bu shi hen qu biaoda na zhong qingxu de dongxi  
They male then NEG be very go express that ZHONG emotion NOM  
DONGXI  
'They boys, don't (like to) express (stuff like) their emotions.'

In the last three examples, *dongxi* seems to be an "empty" head noun. What the speaker intends to indicate by the NP was what the modifying noun refers to (i.e., grammar, theories, emotion, respectively). The "N – de – *dongxi*" modification construction involving *dongxi* as the head noun appears to be a rhetorical strategy in spontaneous communication, through which a lengthened expression is created to highlight the message the speaker intends to get across. The "lengthened form for

important content” strategy is a typical application of the principle of linguistic iconicity to foregrounding.

In summary, the data shows that *dongxi* strongly suggests referential underspecificity. Sometimes its use reflects the intended underspecificity in the referent; however, in some other cases *dongxi* is used instead of a more specific noun because of the spontaneous nature of conversation. In yet some other cases, while the referent can be clearly identified by the modifying expression, the “empty” *dongxi* still serves as the head noun so as to create a lengthened form in order to reflect the content that the speaker wants to foreground. On the other hand, vagueness can also be found in linguistic elements that co-occur with *dongxi*, such as various modifying expressions and measure words. Finally, *dongxi*’s tendency of indicating referential underspecificity is confirmed by the fact that it has a higher co-occurrence rate with general measure words (e.g., *xie* and *zhong*) than with specific ones (e.g., *ge*).

### 3.3 SHI(QING)

We take the traditional perspective and treat *shi* and *shiqing* as the one-syllable and two-syllable variants of the same lexical item. There are altogether 204 tokens of *shi(qing)* in our database (121 *shi* and 83 *shiqing*). However, *shi* is frequently combined with other words to form idiomatic expressions. In the following, we discuss *shi* idiomatic expressions in 3.3.1 and the non-idiomatic uses of *shi(qing)* in 3.3.2.

#### 3.3.1 *Shi* in idiomatic expressions

In our conversation database, 44 tokens of *shi* constitute part of an idiomatic expression, which make up 22% of the total 204 *shi(qing)* tokens, or 36% of the total 121 *shi* tokens. The most frequent expressions are *meishi* (n=23) and *youshi* (n=8), but there are 13 other expressions, such as the interrogative *zemehuishi* ‘what happened?’ and *shi* SOMEONE *de shi* ‘(It) is SOMEONE’s business.’

*Meishi* is in general positive, while *youshi* is in general negative. *Meishi* can express various meanings depending on the speech context, including at least the following: (a) ‘free, not busy/occupied,’ (b) ‘without trouble, survived (some crisis),’ (c) used metalinguistically, the speaker clarifies that (contrary to the interlocutor’s assumption) s/he has nothing to say (at the moment), (d) as an apology acceptance expression, ‘That’s OK,’ and (e) an emphatic expression meaning ‘for no reason.’ The following are examples:

(17) *Meishi* 1 ‘free, not busy/occupied’

這三天麼，反正沒事啊，就找一個時間出去玩，

Zhe san tian ma, fanzheng meishi a, jiu zhao yi ge shijian chuqu wan

This three day PRT anyway NEG-SHI PRT then find one GE time go-out  
play

‘As for these three days, I don’t have anything to do anyway, so I’ll find  
some time to go out and have some fun.’

(18) *Meishi* 2 ‘without trouble, survive (some crisis)’

擦擦藥應該就沒事了吧。

Ca ca yao yinggai jiu meishi le ba

spread spread medicine should then NEG-SHI PRT PRT

‘It (i.e., the wound) should be okay after medicine is applied.’

(19) *Meishi* 3 ‘(the speaker) has nothing to say (at this point)’

B: 我覺得-

wo juede-

I feel

‘I think-’

A: 什麼?

Sheme?

what

‘What?’

B: 沒事

Meishi meishi.

NEG-SHI NEG-SHI

‘No, nothing.’

(20) *Meishi* 4 (as an apology acceptance expression) ‘that’s OK.’

C: 對不起。

Duibuqi.

sorry

‘Sorry.’

B: 沒事。

Meishi

NEG-SHI

‘It’s okay.’

(21) *Meishi* 5: ‘for no reason’

(reasoning that placing the bookcase in the kitchen is strange)

奇怪我幹嘛沒事炒菜的時候還要看到書櫃?

Qiguai wo ganma meishi chao tsai de shihou hai yao kandao shugui?  
 Strange I why NEG-SHI fry vegetable NOM time still want see bookcase  
 ‘Why do I – (for no reason) – have to see the bookcase when I am  
 cooking?’

By contrast, *youshi* can mean (a) ‘busy, occupied,’ or (b) ‘have trouble,’ in different contexts. Consider the following examples:

(22) *Youshi* 1: ‘busy, occupied’

因為那天我有事不能去聽。

yinwei na tian wo youshi bu neng qu ting

because that day I have-SHI NEG can go listen

‘I got something to do on that day, so I did not manage to attend (the lecture).’

(23) *Youshi* 2: ‘have trouble’

(about what the speaker was thinking right after the 921 earthquake)

我想說e2 我家會不會有事啊！

Wo xiang shuo e2 wo jia hui bu hui you shi a

I think say PRT I house will NEG will have-SHI PRT

‘I was thinking that well...will anything bad happen to my house?’

We can see that *meishi* 1 and *youshi* 1, and *meishi* 2 and *youshi* 2, are respectively paired in meaning. In almost all of these idiomatic expressions, but especially transparent in *meishi* 2 and *youshi* 2, the “matter” that is referred to is negatively interpreted. This characteristic of negative polarity is also associated with the fact that the *youshi* 2 reading is possible only in interrogative but not in declarative. On the other hand, *meishi* 3 and *meishi* 4, each being a contextualized reading, appear in highly specified speech acts, i.e., metalinguistic exchange in spontaneous dialogue for the former and apologizing-acceptance for the latter. Finally, *meishi* 5 is certainly a lexicalized expression, in which *shi* no longer refers to “matter” but rather to the more abstract idea of “reason” or “motive.” The situation being referred to as having happened “for no (good) reason” is typically, in the speaker’s view, something negative or undesirable. The *meishi* 5 expression is thus often used together with the rhetorical question form, just as in example (21).

### 3.3.2 *Shi(qing)* in non-idiomatic uses

There are a total of 160 tokens of *shi(qing)* (77 *shi* and 83 *shiqing*) that are used

non-idiomatically in our database.

With or without modifying expressions, about 39% (n=62) of these tokens take a measure word along. While the general measure words (n= 27, 44% of 62) and the specific measure words (n= 35, 56%) are distributed about evenly, the most frequently co-occurring measure words are *jian* (42%, n=26) and *zhong* (27%, n=17). Consider the following example:

- (24) 你講到這個，我就想到一件事，  
 Ni jiang dao zhe ge, wo jiu xiang dao yi jian shiqing  
 You talk to this GE I then think of one JIAN SHIQING  
 ‘The thing you just mentioned reminds me of one thing.’
- (25) 我覺得不會吧沒有聽說過有這種事，  
 wo jue de bu hui ba mei you ting shuo guo you zhe zhong shi  
 I feel NEG will PRT NEG listen say EXP have this ZHONG SHI  
 ‘I don’t think so, never heard of anything like this before.’

With or without MW, almost half of the *shi(qing)* tokens (n= 68, 43%) occur with modifying expressions. This is similar to the case of *dongxi*.

Different from *dongxi*, *shi(qing)* at the object position co-occurs with some specific types of verb frequently: (1) verbs of doing, especially *zuo* ‘do,’ (2) verbs of happening (e.g., *fasheng* ‘occur, happen, take place’), and (3) verbs of speaking (e.g., *jiang* ‘say’).

The *zuo ... shi(qing)* construction shows up in 19% of the data (n=31, out of all 160 non-idiomatic uses). The construction is most prominently realized with the bare form, rendering *zuo shi(qing)* (61%, n=19). *Zuo shi(qing)* can have two readings: (a) ‘have a job, employed,’ and (b) ‘getting work done.’ Consider the following examples:

- (26) *Zuo shi* :1 ‘go to work, have a job, employed’
- B: 你現在是在讀書嗎?  
 Ni xianzai shi zai dushu ma  
 You now be PROG study Q  
 ‘Are you a student now?’
- A: 我現在做事了。  
 Wo xianzai zuo shi le  
 I now do SHI PRT  
 ‘I am working now.’

(27) *Zuo shi* 2: ‘do work, get work done’

那個時候你可能剛好在做事情沒注意呀。

Na ge shihou ni keneng ganghao zai zuo shiqing mei zhuyi ya

That GE time you maybe just PROG do thing NEG notice PRT

‘At the time maybe you were doing something and did not pay attention to (it).’

*Shiqing* also shows a tendency to co-occur with “happening verbs” such as *fasheng* ‘take place, happen’, *yudao* ‘come across, run into,’ *pengdao* ‘come across, run into,’ and *chu* ‘take place, happen.’ Consider the following examples:

(28) *Fasheng* + *shi(qing)*

你怎麼都不問我，我都發生甚麼事。

Ni zenme dou bu wen wo wo dou fasheng shenme shi

You why all NEG ask I I all happen what SHI

‘Why don’t you ask me what happened to me?’

(29) *Yudao* + *shi(qing)*

這次剛好遇到這事情，

Zhe ci ganghao yudao zhe shiqing

this time just encounter this thing

‘It just so happened that we ran into this incident this time.’

The collocation of verbs of speaking with *shi(qing)* is also outstanding (n=19, 12% of all 160 non-idiomatic uses). Verbs of speaking, such as *jiang* ‘say,’ *tan* ‘talk,’ *gaosu* ‘tell,’ and *liao* ‘chat,’ typically co-occur with MW + *shi(qing)*, whereas the bare form *shi(qing)* is not preferred. The most frequent collocation is the *jiang (na)(yi) jian shi(qing)* construction (n=14, 74% of the 19 verb of speaking cases). Consider the following example:

(30) 我再跟你講一件事。

Wo zai gen ni jiang yi jian shi

I again with you tell one JIAN SHI

‘Let me tell you one more thing.’

When verbs of saying are used with *shi(qing)*, *shi(qing)* typically refers to a new topic that the speaker wants to open up and talk about in the upcoming conversation. The *jiang (na) (yi) jian shi(qing)* construction is stabilized with the emergence of this metalinguistic function; its stabilization also explains the frequent collocations

between *jiang* and *shi(qing)* and between *jian* and *shi(qing)*.

In summary, several things stand out about *shi(qing)*. First, *shi* is frequently used in phrasal expressions, each having its idiomatic, and sometimes multiple, meanings. The most frequently seen expressions in interactive discourse are *meishi* and *youshi*, in which the “matter” that is referred to is mostly negatively interpreted. Thus, *meishi* is generally good and *youshi* is generally not good. In its non-idiomatic uses, the “matter” that *shi(qing)* refers to is not always negative and remains neutral. *Shi(qing)* prefers specific measure words (e.g., *jian*) to general ones (e.g., *zhong*), which is different from the case of *dongxi* and suggests that in interactive discourse *shi(qing)* is more likely to refer to specific, individual matters rather than a collection of matters or a generic reference of matters. Also different from the case of *dongxi*, *shi(qing)* at the object position is frequently associated with several types of verbs: verbs of doing (*zuo*), of happening (*fasheng*, etc.), and of speaking (*jiang*, etc.). Frequent collocates include *zuo (...) shi(qing)* and *jiang (na) (yi) jian shi(qing)*. The collocation with verbs of speaking is often used metalinguistically in conversation, signaling discourse transition.

### 3.4 REN

Compared to *dongxi* and *shi(qing)*, *ren* has a much larger token sample (over 500) in our database. Practically speaking, it was difficult to tabulate all the available data manually. Therefore, the numbers provided in the following discussion are round numbers only.

Given the large sample size and the fact that our database is not tagged, it was difficult to code all the *ren* tokens into the bare/MW/modifying categorization. For one thing, the modifying expressions are too numerous and diverse to check using the keyword search method. Nonetheless, a close examination of the data leaves the impression that the “*hen duo + ren*” construction occurs quite frequently. There are about 30 tokens, the majority of which serves as the A or S in the clause. For example:

(31) 可是還是很多人相信中藥啊。

Keshi haishi hen duo ren xiangxin zhongyao a

But still very many person believe Chinese-medicine PRT

‘But there are still lots of people believing in traditional Chinese medicine.’

Among measure words, *ge* far outnumbered other candidates as the most favorite

co-occurring MW. While there are only 20 some tokens of the “... *xie + ren*” combination and about a dozen tokens of the “... *zhong + ren*” combination, there are more than 160 tokens of the “... *ge + ren*” combination.<sup>4</sup> The most striking thing about the “X + *ge + ren*” construction is that it may be used to provide referential information in addition to serving the typical A, S, or O role in a clause: many of the tokens occur after a noun or a pronoun as an emphatic apposition or before the verb as an emphatic adverbial. The X in the construction is prototypically a numeral. When the construction is used as an apposition, a determiner, such as *zhe* ‘this’ or *na* ‘that’, can be part of the X. When the construction serves as an adverbial, no determiner is in X and the X is usually *yi* ‘one,’ forming the expression *yi ge ren* ‘alone.’ As an apposition, the identity (when a determiner is used) and the quantity (suggested by the numeral) of the referent is foregrounded. Consider examples (32) and (33). As an adverbial, the quantity (usually one) of the referent (involved in the state/action) is foregrounded. Consider example (34). Over 20% of the “X + *ge + ren*” tokens are used for this foregrounding function.

(32) 妳這個人真是可怕。

Ni zhe ge ren zhen shi kepa  
 You this GE REN really be incredible  
 ‘You are really incredible.’

(33) 我們三個人爲了刻這個印章在這邊待好久好久，

Women san ge ren wei le ke zhe ge yinzhang zai zhebian dai haojiu haojiu  
 We three GE REN for COMPL carve this GE seal at here stay long long  
 ‘We three stayed here for quite a long time to wait for the seal to be carved.’

(34) 我以爲你一個人去。

Wo yiwei ni yi ge ren qu  
 ‘I thought you went there alone.’

In its bare form, *ren* can appear in either A/S or O position, with no noticeable skewing either way. As an object, there is no particular association between *ren* and any verb, either. However, the pivotal construction involving the presentative *you*, “*you + ren + ... + (S)V*,” is the most frequently seen collocation when *ren* shows up in its bare form (43%, about 90 out of the 210 bare *ren* tokens). The construction is found in positive declarative, negative declarative, and interrogative as in the

<sup>4</sup> Note that constructions in which the modifying element occurring between the MW and the noun, i.e., “... *ge + MODIFIER + ren*,” “...*xie + MODIFIER + ren*,” and “... *zhong + MODIFIER + ren*,” are not tabulated here for statistics, although we assume the same skewing collocation pattern exists among the three.

following examples:

- (35) 然後那就有人這樣對著鏡頭講，  
 Ranhou na jiu you ren zhe yang dui zhe jingtou jiang  
 Afterward then then have REN this way to DUR lens talk  
 ‘Then there was somebody talking to the camera lens like this,’
- (36) 沒有人要理她。  
 Meiyou ren yao li ta  
 NEG REN want pay-attention-to 3S  
 ‘No one wants to talk to her.’
- (37) 有沒有人遲到啊？  
 you mei you ren chidao a  
 have NEG have REN late PRT  
 ‘Was there anyone late?’

The previous examples show that this construction is useful in talking about discourse participants whose specific identity is (at this point) peripheral to the development of the discourse – thus the bare form of *ren* to indicate low manipulability (Hopper & Thompson, 1984). This referential underspecificity is a consequence of the information flow in on-line speech production. Due to the “one new idea constraint” principle, which limits an intonation unit in conversation to contain no more than one piece of new information (Chafe, 1994), when the message to get across concerns a new discourse participant’s action or state, this person’s referential information tends to remain at the minimum, thus the bare form (rather than alternatives with more nominal information) is preferred in the pivotal construction.

While the pivotal construction is instantiated mainly with the bare *ren*, there is still a variation containing *xie*, rendering “*you + (yi) xie + ren + ... + (S)V*” (about 10 tokens). Since *xie* suggests indeterminate quantity, the identity of the referent remains underspecified. For example:

- (38) 可是有些人錄起來好聽啊，  
 keshi you xie ren lu qilai hao ting a  
 but have XIE REN record COMPL good listen PRT  
 ‘But some people’s voices sound very good on tape.’

On the other hand, some token examples of the “*you + ren + ... + (S)V*” construction are used for the listing function. Again in this function the identity about the people referred to remains underspecified. Consider the following examples:

(39) (about the various purposes people visit Japan)

- a. 因為我們去日本旅遊的目的很多啦。  
 yinwei women qu Riben luyou mudi hen duo la  
 because we go Japan travel purpose very many PRT  
 ‘It is because that we travel in Japan for various purposes.’
- b. 有人可能是去看風景，  
 you ren keneng shi qu kan fengjing  
 have REN maybe be go look scenery  
 ‘Some people probably go there for sight seeing.’
- c. 有人去看偶像，  
 you ren qu kan ousiang  
 have REN go see idol  
 ‘Some people go there to see their idols.’
- d. 有人去 shopping，  
 you ren qu shopping  
 have REN go shopping  
 ‘Some people go there for shopping.’
- e. 有人是純感受這樣子，  
 you ren shi chun ganshou zheyangzi  
 have REN be purely feel this-way  
 ‘And still some people go there just to feel it (i.e., its atmosphere).’

Of course, this listing function can also be expressed by the variant “*you + (yi) xie + ren + ... + (S)V*,” or by another expression, *you de ren* ‘some people’ (about 20 tokens in our database). Consider the following example for the latter expression:

- (40) a. 她們有的人去染髮啊，  
 Tamen youderen qu ranfa a  
 They some go dye-hair PRT  
 ‘Some of them got their hair dyed.’
- b. 然後有的人是換了一個完全新的髮型，  
 Ranhou youderen shi huan le yi ge wanquan xin de faxing  
 Afterward some be change COMPL one GE completely new NOM  
 hair-style  
 ‘and some of them had a brand new hairstyle.’

In summary, while *ren* is in general evenly distributed with regard to syntactic

positions, there are still some noticeable collocations. It appears that *ren* is used for both referential specificity and underspecificity. The evidence for the former is the fact that *ge* is its favorite collocating measure word. However, while “*X gen ren*” is commonly seen in all major syntactic positions (A/S and O), a good chunk of the “*X ge ren*” construction is used as an apposition or an adverbial, both to foreground information about the referent. On the other hand, *ren* is also commonly seen in bare form, suggesting referential underspecificity. When in bare form, it combines with the presentative *you* to form a pivotal construction, “*you + ren + ... + (S)V*,” which is widely used in conversation to introduce a discourse participant whose identity is not part of the prominent information.

#### 4. Conclusion

In this paper, we have studied how three Chinese general nouns, *ren*, *shi(qing)*, and *dongxi*, are used in spoken Mandarin. The analysis shows that each lexical item displays different tendencies with regard to referential specificity. *Ren* appears to be the least “marked” in this respect: whereas its frequent collocation with *ge* suggests referential specificity, its equally frequently seen bare form indicates referential underspecificity. Among the three general nouns, *ren* is quite unique in being frequently paired up with the presentative/existential *you* to form a pivotal construction to introduce a new human participant and his/her action/state into the discourse. The appositional and adverbial uses involving the “*X gen ren*” construction, which foregrounds properties of the referent (quantity and/or identity), is also conspicuous in spoken discourse. On the other hand, both *dongxi* and *shi(qing)* designate non-human referents, but they differ in the extent to which they indicate referential specificity. *Dongxi*, designating either concrete objects or abstract ideas, tends to provide vague, underspecified identification of the referent. Measure words or modifying expressions co-occurring with it also tend to convey vagueness. In contrast, *shi(qing)* tends to refer to specific events or matters and is more likely to occur with specific measure words.

In this study we are confronted with the abundance of syntagmatic combinations of linguistic elements that go beyond NP. For example, there are verb-noun collocations (e.g., *zuo + shiqing*), constructions of various structural complexities (with discourse functions that are undergoing conventionalization) (e.g., “*jiang ... (na) (yi) jian shi(qing)*”), and phrasal expressions with idiomatic meanings (e.g., *meishi*). In fact, as repeatedly pointed out in the corpus literature, the high frequency of lexical items may sometimes be due to their frequent recurrence in set phrases or collocations (Sinclair, 1999; Stubbs, 2002). In this study, the recurrent

appearance of *shi* in set phrases such as *meishi* and *youshi*, which are themselves frequently used due to their multiple conventionalized meanings, partially accounts for its overall high frequency in spoken discourse.

In this study, we have also observed the impact of conversation as a type of interactive discourse onto the ways speakers use their language and the ways meaning is constructed, reinforced, and conventionalized in language use. The spontaneous nature of conversation may have motivated some of the profuse uses of general nouns, such as the case in which *dongxi* merely functions as an “empty-head”, as a strategy to create a lengthened expression to iconically highlight what the speaker wants to foreground. Next, the metalinguistic communication between interlocutors in conversation may have contributed to the “stabilization” of the frequent collocation of *shi* with specific measure words and verbs of speaking into a construction, e.g., *jiang (na) (yi) jian shi(qing)*. Finally, the pivotal construction involving the bare *ren* (sometimes also with *xie*) and the presentative/existential *you* is also a consequence of the “one new idea constraint” in on-line speech production. When the newly introduced discourse participant’s action or state is elaborated in the pivotal structure, *ren*, the bare form (rather than other more informative NP alternatives), is preferred in the presentative/existential structure.

This study, which is limited in its scope, methodology, and database, is only the first step toward our quest into general nouns and, more generally, into the relationship between lexis, grammar and discourse. For future research, more general nouns should be covered to broaden the scope of the investigation. Second, the data to be examined is preferably tagged so that large quantities of samples can be computationally handled. Finally, text types other than conversation should be checked in order to identify more genre factors in meaning construction.

## References

- Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad, & Edward Finegan. 1999. *Longman Grammar of Spoken and Written English*. London: Longman.
- Biq, Yung-O. 1990. Question words as hedges in conversational Chinese: A Q and R exercise. *Pragmatics and Language Learning*, Monograph Series vol. 1, ed. by Lawrence B. Bouton and Yamuna Kachru, 141-157. Urbana-Champaign: University of Illinois.
- Biq, Yung-O. 2004. From collocation to idiomatic expression: The grammaticalization of *hao* phrases/constructions in Mandarin Chinese. *Journal of Chinese Language and Computing*, 14.2. Special Issue: Corpora, Language Use, and Grammar.
- Biq, Yung-O. To appear. Construction, reanalysis, and stance: ‘V yi ge N’ and variations in Mandarin Chinese. *Journal of Pragmatics*.
- Bybee, Joan & Paul Hopper (ed.). 2001. *Frequency and the Emergence of Linguistic Structure*. Amsterdam: Benjamins.
- Chafe, Wallace. 1994. *Discourse, Consciousness, and Time*. Chicago: University of Chicago Press.
- Cowie, A.P. (ed.). 1998. *Phraseology: Theory, Analysis, and Applications*. Oxford: Clarendon Press.
- Erman, Britt, & Beatrice Warren. 2000. The idiom principle and the open choice principle. *Text* 20.1: 29-62.
- Fang, Mei (方梅). 2002. <指示詞“這”和“那”在北京話中的語法化> [The grammaticalization of the demonstratives *zhe* and *na* in Beijing Mandarin] , 《中國語文》 [*Zhongguo Yuwe*] 2002.4: 343-356。
- Francis, Gill. 1994. Labelling discourse: an aspect of nominal-group lexical cohesion. *Advances in Written Text Analysis*, ed. by Malcolm Coulthard, 83-101. London: Routledge.
- Halliday, M.A.K. & Ruqaiya Hasan. 1976. *Cohesion in English*. London: Longman.
- Hopper, Paul. 1988. Emergent grammar and the a priori grammar constraint. *Linguistics in Context*, ed. by Deborah Tannen, 117-134. Norwood, NJ: Ablex.
- Hopper, Paul. 1998. Emergent Grammar. *The New Psychology of Language*, ed. by Michael Tomasello, 155-175. Hillsdale, NJ: L. Erlbaum.
- Hopper, Paul & Sandra A. Thompson. 1984. The discourse basis for lexical categories in universal grammar. *Language* 60: 703-51.
- Huang, Chu-Ren (ed.). 2003. *Language and Linguistics* 4.3., Special Issue: Chinese Lexical Semantics.
- Huang, Shuanfan. 1999. The emergence of a grammatical category *definite article* in

- spoken Chinese. *Journal of Pragmatics* 31: 77-94.
- Huang, Shuangfan. 2000. The story of heads and tails – on a sequentially sensitive lexicon. *Language and Linguistics* 1.2: 79-107.
- Hunston, Susan & Gill Francis. 2000. *Pattern Grammar: A Corpus-driven Approach to the Lexical Grammar of English*. Amsterdam: Benjamins.
- Lenk, Uta. 2000. Stabilized expressions in spoken discourse: Worth our time? *Corpus Linguistics and Linguistic Theory* (ICAME 20), ed. by Christian Mair & Marianne Hundt, 187-200. Amsterdam: Rodopi.
- Liu, Mei-chun. 2001. *Mandarin Verbal Semantics: A Corpus-based Approach*. Taipei: Crane.
- Mahlberg, Michaela. 2003. The textlinguistic dimension of corpus linguistics: The support function of English general nouns and its theoretical implications. *International Journal of Corpus Linguistics* 8.1: 97-108.
- Moon, Rosamund. 1998. *Fixed Expressions and Idioms in English: A Corpus-based Approach*. Oxford: Clarendon Press.
- Ochs, Elinor, Emanuel A. Schegloff, & Sandra A. Thompson (ed.). 1996. *Interaction and Grammar*. Cambridge: Cambridge University Press.
- Partington, Alan. 1998. *Patterns and Meanings: Using Corpora for English Language Research and Teaching*. Amsterdam: Benjamins.
- Sinclair, John. 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Sinclair, John. 1999. A way with common words. *Out of Corpora*, ed. by Hilde Hasselgard & Sigrun Oksefjell, 157-179. Amsterdam: Rodopi.
- Stubbs, Michael. 2001. *Words and Phrases: Corpus Studies of Lexical Semantics*. Oxford: Blackwell.
- Stubbs, Michael. 2002. Two quantitative methods of studying phraseology in English. *International Journal of Corpus Linguistics* 7.2: 215-244.
- Tao, Hongyin (陶紅印). 2000. <從‘吃’看動詞論元結構的動態特徵> [‘Eating’ and emergent argument structure], 《語言研究》 [*Language Research*] 2000.3: 21-38。
- Tao, Hongyin (陶紅印). 2003a. <從語音語法和話語特徵看“知道”格式在談話中的演化> [Phonological, grammatical, and discourse evidence for the emergence of *zhidao* constructions], 《中國語文》 [*Zhongguo Yuwen*] 2003: 291- 302。
- Tao, Hongyin. 2003b. Toward an emergent view of lexical semantics. *Language and Linguistics* 4.4: 837-856.
- Thompson, Sandra A. 1988. A discourse approach to the cross-linguistic category ‘adjective.’ *Explaining Language Universals*, ed. by John Hawkins, 167-185. Oxford: Blackwell.

Zhang, Bojiang, & Mei Fang (張伯江&方梅). 1996. 《漢語功能語法研究》 [*A Study of Chinese Functional Grammar*]. 南昌：江西教育出版社 [Nanchang: Jiangxi Jiaoyu Chubanshe].

Zou, Shaohua (鄒韶華). 2001. 《語用頻率效應研究》 [*A Study of Frequency Effects of Language in Use*]. 北京：商務 [Beijing: Shangwu].

[Received 26 February 2004; revised 30 April 2004; accepted 4 May 2004]

Department of English  
National Taiwan Normal University  
Taipei, TAIWAN  
ybiq@cc.ntnu.edu.tw

## 人、事、物：華語口語中的概述名詞

畢永峨

國立台灣師範大學

本文從語料庫語言學的角度研究當代台灣華語口語中「人」、「事(情)」、「東西」三個概述名詞的使用。我們專門探討這三個詞與其他語言成分共同出現的情形。我們的調查顯示這三個詞對於指涉的區辨程度各有不同傾向。每個詞在與其他語言成分共同組成習語或半固定句式的程度上也不同。這三個詞的使用也呈現出語言互動如何驅動並強化意義建構的制式化。

關鍵詞：概述名詞、語料庫為本的詞彙語法研究、共現、習語、句式