Collocation, semantic prosody and near synonymy:

A cross-linguistic perspective

Abstract: This paper explores the collocational behaviour and semantic prosody of near synonyms from a cross-linguistic perspective. The importance of these concepts to language learning is well recognized. Yet while collocation and semantic prosody have recently attracted much interest from researchers studying the English language, there has been little work done on collocation and semantic prosody on languages other than English. Still less work has been undertaken contrasting the collocational behaviour and semantic prosody of near synonyms in different languages. In this paper, we undertake a cross-linguistic analysis of collocation, semantic prosody and near synonymy, drawing upon data from English and Chinese (*pu3tong1hua4*). The implications of the findings for language learning are also discussed.

1. Introduction

This paper explores the collocational behaviour and semantic prosody of near synonyms from a cross-linguistic perspective. The importance of these concepts to language learning is well recognized. Yet while collocation and semantic prosody have recently attracted much interest from researchers studying the English language, there has been little work done on collocation and semantic prosody on languages other than English. Still less work has been undertaken contrasting the collocational behaviour and semantic prosody of near synonyms in different languages. The

exceptions to this include Sardinha (2000) and Tognini-Bonelli (2001: 131-156), who contrast semantic prosodies in European languages such as English vs. Portuguese and English vs. Italian. In this paper, we will use the same methodology to undertake a cross-linguistic analysis of collocation and semantic prosody of near synonyms, drawing upon data from two distinctly different languages, English and standard Chinese (*pu3tong1hua4*).[1]

In working with English and Chinese we decided that we needed to carefully select corpus data to minimize the role that corpus choice had to play in determining our research findings. We knew that our approach should be corpus-based as previous studies have shown that a speaker's intuition is usually an unreliable guide to patterns of collocation and that intuition is an even poorer guide to semantic prosody. Yet in choosing corpora for our study we also wanted to use data that formed a good basis for a contrastive study, as we wished to control for features such as genre when undertaking our cross-linguistic comparison. Consequently, the principal corpora we use in this paper are the Freiburg-LOB Corpus of British English (i.e. FLOB, see Hundt, Sand and Siemund 1998), the Freiburg-Brown Corpus of American English (i.e. Frown, see Hundt, Sand and Skandera 1999), and the Lancaster Corpus of Mandarin Chinese (i.e. LCMC; see McEnery, Xiao and Mo 2003). Each of these corpora contains approximately one million words of samples collected from fifteen written text categories published around 1991. The FLOB/Frown and LCMC corpora are, as far as is practically possible, comparable corpora suitable for contrastive language research. However, there were points in our research when these corpora were not large enough to provide a reliable basis for quantification. On such occasions a supplementary group of data was also used, which includes two comparable corpora, namely the Lancaster-Oslo/Bergen Corpus of British English (i.e. LOB, see Johansson,

Leech and Goodluck 1978) and the Brown University Corpus of American English (i.e. Brown, see Francis and Kučera 1964), and the People's Daily (2000) Corpus for Chinese (PDC2000).[2] The LOB and Brown corpora are 1-million-word corpora sampled from the same fifteen text categories used in FLOB, Frown and LCMC, but the texts were produced in 1961. The PDC2000 corpus covers one year's newspaper texts published by the People's Daily in 2000, totalling approximately fifteen million words. For very rare items in English, we also used the British National Corpus (i.e. BNC, World Edition) to extract significant collocates. As the supplementary corpora are not comparable either in sampling period or coverage, we clearly indicate where they are used in this paper. These corpora were only used to add further weight to observations made in small comparable corpora.

In this paper, collocation, semantic prosody and near synonymy are explored contrastively with the aim of addressing two important questions: i.) does Chinese exhibit semantic prosody and semantic preference as English does? ii.) how different (or similar) are the collocational behaviour and semantic prosody of lexical items with similar denotational meanings (i.e. near synonyms) in unrelated languages such as English and Chinese? In addition, in the conclusion of this paper, we will also consider the implications of this contrastive analysis for language learning.

To make the analysis comparable, the exploration will be undertaken through three case studies of near synonyms in English and their close translation equivalents in Chinese. These translation equivalents were identified using a bilingual dictionary, independently of the observed collocational behaviour or semantic prosody in our corpus data. When used such dictionary-based translation equivalents were checked by us in context to exclude irrelevant data. For example, the dictionary shows the Chinese equivalent of *result* is *jie2guo3* (结果), which can also be used as a verb

meaning 'to yield fruits' literally, or 'to produce a result' and, metaphorically, 'to kill or finish off somebody'. These uses were excluded from our analysis, as they fell outside the scope of our study. Before proceeding to undertake the contrastive analysis, however, we will first review previous research on collocation, semantic prosody and near synonymy.

2. Collocation, semantic prosody and near synonymy: a review

Collocation has been studied for at least five decades. The word *collocation* was first used as a technical term by Firth (1957) when he said 'I propose to bring forward as a technical term, meaning by *collocation*, and apply the test of *collocability*' (Firth 1957: 194). According to Firth (1968: 181), 'collocations of a given word are statements of the habitual or customary places of that word'. Firth's notion of collocation is essentially quantitative (cf. Krishnamurthy 2000: 32). The statistical approach to collocation is accepted by many corpus linguists including, for example, Halliday (1966: 159), Greenbaum (1974: 82), Sinclair (1991), Hoey (1991), Stubbs (1995), Partington (1998), McEnery and Wilson (2001), and Hunston (2002). All of these linguists follow Firth in that they argue that collocation refers to the characteristic co-occurrence of patterns of words. One assumes that Greenbaum's (1974: 82) definition of collocation – 'a frequent co-occurrence of two lexical items in the language' – only refers to statistically significant collocation. While Greenbaum's definition does not tell us how frequent the co-occurrence of two lexical items should be to be considered as a collocation, Hoey (1991: 6-7) uses the term *collocation* only if a lexical item appears with other items 'with greater than random probability in its (textual) context'.

There are a number of statistical tests used to measure collocational strength, e.g. the MI (mutual information), *z*, *t*, log-likelihood, log-log and MI3 scores. While it is well known that the MI scores may unduly overvalue infrequent words, we chose to

use the MI measure in this paper because it is built into the corpus tools we used, WordSmith (for English data) and Xaira (for Chinese data). Both tools allow users to set the minimum co-occurrence frequency of an item to be considered as a collocate of a given node word so that the drawback of the MI measure, as noted above, can be partly offset. Given the size of the comparable corpora used, we set the minimum co-occurrence frequency to 3. Within a 4-4 window span, items which have a minimum co-occurrence frequency of 3 as a collocate of a given node word and a minimum MI score of 3 are considered to be collocates of a node word. When using additional data from the BNC and PDC2000 corpora, the minimum co-occurrence frequency was set at 20. As we will see from the collocates extracted in section 3, these adjustments have allowed us to use the MI score safely.

Shifting from form to meaning, Stubbs (2002: 225) observes that 'there are always semantic relations between node and collocates, and among the collocates themselves'. The collocational meaning arising from the interaction between a given node and its typical collocates might be referred to as *semantic prosody*, 'a form of meaning which is established through the proximity of a consistent series of collocates' (Louw 2000: 57). Both individual words and phrases can have semantic prosodies (cf. Schmitt and Carter 2004: 7). The primary function of semantic prosody is to express speaker/writer attitude or evaluation (Louw 2000: 58). Semantic prosodies are typically negative, with relatively few of them bearing an affectively positive meaning. However, a speaker/writer can also violate a semantic prosody condition to achieve some effect in the hearer – for example irony, insincerity or humour can be explained by identifying violations of semantic prosody (Louw 1993: 173).

It would appear, from the literature published on semantic prosody, that it is at least as inaccessible to a speaker's conscious introspection as collocation is (cf. Louw 1993: 173; Partington 1998: 68; Hunston 2002: 142). Yet as the size of corpora has grown, and tools for extracting semantic prosodies have been developed, semantic prosodies have been addressed much more frequently by linguists, as exemplified in Table 1.[3]

[TABLE 1 NEAR HERE]

It might be argued that the negative (or less frequently positive) semantic prosody that belongs to an item is the result of the interplay between the item and its typical collocates. On the one hand, the item does not appear to have an affective meaning until it is in the context of its typical collocates. On the other hand, if a word has typical collocates with an affective meaning, it may take on that affective meaning even when used with atypical collocates. As the Chinese saying goes, 'he who stays near vermilion gets stained red, and he who stays near ink gets stained black' – one takes on the colour of one's company – the consequence of a word frequently keeping 'bad company' is that the use of the word alone may become enough to indicate something unfavourable (cf. Partington 1998: 67).

In Stubbs' (2002: 225) comment cited above, the meaning arising from the common semantic features of the collocates of a given node word can be referred to *semantic preference*, which is defined 'by a lexical set of frequently occurring collocates [sharing] some semantic feature' (*ibid*: 449). For example, Stubbs (2001b: 65) observes that *large* typically collocates with items from the same semantic set indicating 'quantities and sizes' (e.g. *number(s)*, *scale*, *part*, *quantities*, *amount(s)*) while Partington (2004: 148) notes that 'absence/change of state' is a common feature of the collocates of maximizers such as *utterly*, *totally*, *completely* and *entirely*.

Semantic preference and semantic prosody are two distinct yet interdependent collocational meanings. According to Sinclair (1996, 1998) and Stubbs (2001b), semantic prosody is a further level of abstraction of the relationship between lexical units: collocation (the relationship between a node and individual words), colligation (the relationship between a node and grammatical categories), semantic preference (semantic sets of collocates) and semantic prosody (affective meanings of a given node with its typical collocates). Partington (2004: 151) notes that semantic preference and semantic prosody have different operating scopes: the former relates the node item to another item from a particular semantic set whereas the latter can affect wider stretches of text. Semantic preference can be viewed as a feature of the collocates while semantic prosody is a feature of the node word. On the other hand, the two also interact. While semantic prosody 'dictates the general environment which constrains the preferential choices of the node item', semantic preference 'contributes powerfully' to building semantic prosody (Partington 2004: 151).

There are different opinions regarding whether or not semantic prosody is a type of connotative meaning. Partington (1998: 68), Stubbs (2001a: 449) and Hunston (2002: 142) appear to take it for granted that semantic prosody is connotational, while Louw (2000: 49-50) explicitly argues that 'semantic prosodies are not merely connotational' as 'the force behind SPs [semantic prosodies] is more strongly collocational than the schematic aspects of connotation'. In our view, connotation can be collocational or non-collocational whereas semantic prosody can only be collocational.

In this paper, the positive, neutral and negative semantic prosodies correspond to Partington's (2004) favourable, neutral and unfavourable prosodies. We evaluated each case in context. A pleasant or favourable affective meaning was labelled as

positive while an unpleasant or unfavourable affective meaning was judged as negative. When what was happening was completely neutral, or the context provided no evidence of any semantic prosody, the instance was labelled as neutral. Note that in the collocate lists presented in section 3, items with an unfavourable affective meaning are underlined and those with a favourable affective meaning are in bold.

In this paper, by near synonyms we mean lexical pairs that have very similar cognitive or denotational meanings, but which may differ in collocational or prosodic behaviour (c.f. also Partington 1998: 77). As such, synonymous words are not collocationally interchangeable (cf. Conzett 1997: 70-87; Tognini-Bonelli 2001: 34). For example, Greenbaum (1974: 81) noted that synonyms may 'be separated collocationally because of restrictions to a language variety or style', as shown in his examples *to cashier an army officer* vs. *to expel a school child*. Halliday (1976: 73) observed that tea is typically described as *strong* rather than *powerful* whereas a car is more likely to be described as *powerful* than *strong*, even though the two modifiers share similar denotational meanings. Similarly, while *weak* and *feeble* have similar cognitive meanings, native speakers of English prefer to say *weak tea* rather than *feeble tea* (cf. Mackin 1978: 150). In addition to different collocational behaviour, near synonyms can also differ in semantic prosodies, e.g. *fickle* is negative whereas *flexible* is positive (see Tognini-Bonelli 2001: 18-24). Finally, while it is important to note that, in addition to the lexical level, near synonyms at the morphological level also demonstrate different collocational behaviour (e.g. cf. Greenbaum 1974: 81-82), in this paper, we will only consider synonyms at the word/phrase level.

3. Collocation, semantic prosody and near synonymy in English and Chinese

Previous research on collocation and semantic prosody has focused on English. Other languages and contrastive studies have largely been overlooked. In order to begin to

address this oversight, in this section, we will contrast the collocational behaviour of near synonyms and their semantic prosodies in English and Chinese. Specifically, the following groups of near synonyms in English and Chinese will be examined: the *consequence* group, the CAUSE group, and the *price/cost* group. These words are selected because they have been studied in English and we want to move from the established patterns of English to investigate the patterns occurring in Chinese by examining the collocational behaviour and semantic prosody of the close equivalents of the words in question in Chinese. For each group, collocation and semantic prosody in English are discussed first, followed by a contrastive analysis of the Chinese data.

3.1 The *consequence* group

In English, there are a number of words that mean 'anything that is due to something already done', e.g. *result*, *outcome*, *consequence* and *aftermath*. While *aftermath* is sometimes labelled as 'unpleasant', most dictionaries simply ignore the different semantic prosodies of these near synonyms with no further distinction (see discussion below). Nevertheless, such definitions are arguably inadequate and misleading. In fact, the affective meanings of near synonyms are closely associated with word forms, contexts and genres. Table 2 shows the distribution of *consequence* across meaning categories in the FLOB/Frown corpora. It is clear from the table that while fixed expressions such as *as a consequence* and *in consequence (of)* can be negative, neutral or positive, depending upon their contexts, *consequence* and *consequences* show a strong tendency towards a negative semantic prosody. The plural form *consequences* is even more likely to be used negatively. When it is used neutrally, *consequence(s)* often means 'importance', or is typically used in contrast with *antecedent/precedence* or *conditional*.

[TABLE 2 NEAR HERE]

As with the neutral use of *consequence(s), consequent(ly)* occurs most frequently in academic prose as a formal substitute of *as a result*. In FLOB and Frown, significant collocates of *consequences* include (ranked by co-occurring frequency):

- Nature: **important**, <u>adverse</u>
- Affected target: *social, financial, economic, ethical, moral, individual, public*
- Action: HAVE, *(there) BE*, ACCEPT

Of the collocates indicating the nature of consequences, *important* is positive while *adverse* is negative. Interestingly, all instances of *important consequences* follow HAVE or *there BE*. In FLOB/Frown we observed that *important consequences* tend to collocate with *HAVE/there BE* to denote a positive pattern meaning. It appears that the two patterns are associated with the different meanings of *consequence*: 'importance' and 'result'. This observation is confirmed in a larger corpus. Of the 67 instances of *important consequences* in the BNC, 54 occurrences follow HAVE and one instance follows *there BE*. All 54 examples are positive. The other 12 cases may either be positive or neutral. When *consequences* means 'results', its typical collocates refer to something unpleasant. In the BNC, for example, collocates indicating the nature of consequences (with a co-occurrence frequency of 20 or above) include, in the order of their co-occurring frequencies, <u>serious</u>, **important**, <u>disastrous</u>, <u>adverse</u>, <u>dire</u>, *far-reaching*, <u>damaging</u>, <u>negative</u>, *profound*, <u>unintended</u>, *major*, <u>unfortunate</u>, <u>tragic</u>, <u>fatal</u>, *new*, <u>severe</u> and **significant**. All of the underlined items express an unfavourable affective meaning.

When they are modified by collocates indicating an affected target, *consequences* are typically negative. As such, actions associated with them normally include *accept, alleviate, avoid, face, meet, minimise, offset, (be) responsible, (take) responsibility,*

suffer and *sustain*.[4] *Consequences* sometimes collocates with verbs such as *REAP*, as in (1):[5]

(1)   These officials generally attributed their problems to: […] Some critics charged, though, that states were reaping the consequences of profligate spending during the growth years of 1984-1989. (Frown: H)

REAP typically collocates in its literal meaning with names of crops and *harvest*, or metaphorically with words with a positive meaning such as *benefit(s)* and *rewards* (the three significant collocates are from the BNC). It seems that the apparently paradoxical combination of *REAP* and *consequences* in this example carries the implication that 'you reap as you sow': the officials were suffering as a result of their own profligate spending.

In comparison with *consequence(s)*, *aftermath* displays an even more pronounced tendency towards the negative pole of the semantic continuum. In FLOB and Frown, 14 occurrences of *aftermath* were found, mostly in the expression *in the aftermath of*. There is only one significant collocate indicating what causes the state of affairs referred to by *aftermath*. It is *war*. As the low frequency may result in unreliable quantification, we consulted the BNC, which provides 687 instances of *aftermath*. Significant collocates in the BNC typically include (with a co-occurrence frequency of 20 or above) <u>*war(s)*</u>, *world* (as in *World War I*) and *Gulf* (as in the *Gulf War*). If we move down the list to include collocates with a co-occurrence frequency of 10 or above, we will find *first* (as in the *First World War*), <u>*death*</u>, <u>*coup*</u>, *second* (as in the *Second World War*), *election, events* and *revolution*. It is clear that most of these words are negative in their contexts.

Further away from the negative pole of the semantic continuum are *result* and *outcome*. *Result* is significantly more common than *outcome* (with 677 and 86

occurrences respectively in FLOB/Frown). It appears that both words are associated with a favourable affective meaning, e.g. *a good result, a great result, an excellent result, a brilliant result, a successful outcome* (cf. Hoey 2003), as reflected by their significant collocates which indicate the nature of a result or outcome (ranked by co-occurring frequency):

- Result: **better**, *different, early, end, final, similar, direct, empirical, likely, experimental*, **good**, *negative*, **desired**

- Outcome: *likely*, **positive**, **successful**

It is of interest to note that *negative* appears on the collocation list of *result*. A close examination of the concordances shows that in all of the three instances *negative* should be interpreted in a mathematical or medical sense, which has no impact upon affective meaning.[6]

While the discussion above shows that the four near synonyms can be arranged, from positive to negative, on a semantic continuum as follows: *outcome/result, consequence* and *aftermath*, dictionaries and thesauri (both learners' and non-learners') are not terribly helpful in distinguishing the different semantic prosodies of near synonyms (cf. Partington 1998: 69-72; Tognini-Bonelli 2001: 25), because they emphasize the denotational meaning of these synonymous words, not their usage. For example, the *Longman Synonym Dictionary* (1986: 33, 217) simply lists *result, effect, event, outcome, issue, upshot* and *aftermath* etc as the synonyms of *consequence* while giving *result, consequences, end result* and *outcome* etc as the synonyms of *aftermath*. The new edition of the *Longman Dictionary of English Language and Culture* is much better in this respect. While it also defines *consequence* neutrally as 'something that follows from an action or set of conditions; result', all of the examples it provides are the negative use of the word (1999: 272). The *Merriam-*

Webster Thesaurus (1991) and the *Oxford Thesaurus* (1991) also overlook the different semantic prosodies of *result* and *consequence* and, in Tognini-Bonelli's (2000: 224) words, 'end up showing a certain circularity'. Only two of the dictionaries that we examined provide an adequate account of the use of *consequence*. The *Cassell Guide to Related Words* (1994: 451-452) gives a clear explanation of the subtle distinction between *result*, *consequence*, *outcome* and *denouement*, according to which '*consequence* may refer to simple causation in a neutral way […] More often, however, *consequence* suggests a negative result or at least the negative concomitant of an otherwise desirable effect'. *Collins English Dictionary* (5th edition, 2000: 340) draws a distinction between the neutral ('a result or effect of some previous occurrence') and negative ('an unpleasant result') use of *consequence* and gives an example of its negative use (*take the consequences*).

Shifting to consider these words in contrast, the Chinese equivalent of *result*/*outcome* is *jie2guo3* (结果) while the equivalent for *consequence*/*aftermath* is *hou4guo3* (后果). In addition, there are a number of obviously positive synonyms such as *cheng2guo3* (成果) 'achievement' and *shuo4guo3* (硕果) 'great achievement', and negative synonyms including *ku3guo3* (苦果) 'a bitter pill to swallow' and *e4guo3* (恶果) 'evil consequence'. There are 240 instances of *jie2guo3* (结果) in the LCMC corpus, which are distributed across different meaning categories as follows: positive 33, neutral 129 and negative 78. Significant collocates of *jie2guo3* (结果) include:

- Modifiers: *da4xuan3* 'general election', *bi4ran2* 'inevitable', *shi4yan4* 'experiment', *diao4cha2* 'investigation', *ke3neng2* 'possible', *jing1ji4* 'economic', **hao3** 'good'

- Actions: *biao3ming2* 'show', *zao4cheng2* 'cause', *zeng1jia1* 'increase', *chan3sheng1* 'give rise to; arise', *you3* 'have'

There are both similarities and differences in the distribution of *result* and its Chinese equivalent *jie2guo3* (结果) across meaning categories. On the one hand, like its English equivalents *result* and *outcome*, *jie2guo3* (结果) typically does not express a negative affective meaning. The semantic prosody of *jie2guo3* (结果) is dependent upon its collocates. For example, when it collocates with *zao4cheng2* (造成), it indicates an unfavourable result; conversely, when it collocates with *chan3sheng1* (产生), the result is evaluated favourably (see section 3.2 for further discussion of *zao4cheng2* (造成) and *chan3sheng1* (产生)). The neutral use of *jie2guo3* (结果) was mainly found in academic prose. As there are inherently positive synonyms in Chinese (e.g. *shuo4guo3* (硕果) and *cheng2guo3* (成果)), as noted above, *jie2guo3* (结果) is less frequently used than *result* to indicate a positive semantic prosody.

In relation to *jie2guo3* (结果), *hou4guo3* (后果) is typically negative, though it can be used neutrally, because in some instances there is no evidence of semantic prosody in its context. Of the 22 occurrences of *hou4guo3* (后果) in the LCMC corpus, 19 are used negatively, with the remaining three being neutral. The only significant collocate of *hou4guo3* (后果) in LCMC is *yan2zhong4* 'serious, grave'. When the consequences are specified, they typically refer to undesirable situations such as *increasingly intensifying contradictions*, *goods piling up in stock* and *inflation*. When the consequences are not made clear, there are usually modifiers expressing value judgements or indicating the nature of the consequences. These modifiers normally share a negative preference, including, for example, *yan2zhong4* 'serious, grave', *bu4kan1she4xiang3* 'too ghastly to contemplate', *bu4ke3wan3hui2*

'irretrievable', *bu4liang2* 'adverse', *xiao1ji2* 'negative', *nan2ce4* 'dubious' and *bu4yan2zi4yu4* 'self-evident'. In fact, *hou4guo3* (后果) keeps bad company so frequently that simply using this word alone is usually sufficient to indicate some unfavourable result, as in (2):

(2) a. *ni3 xiang3xiang3 na4 hui4 you3 zen3yang4 de hou4guo3*

   (LCMC: G)

   you think-think that will have what GEN consequence

   'Just imagine what consequences will result.'

  b. *hng-hng, na4 hou4guo3, qing3 xian1sheng1 zi4ji3 hao3sheng1*

   Humph then consequence please mister self carefully

   xiang3xiang3 ba (LCMC: N)

   think-think PRT

   'Humph! Then you must think carefully of the consequences.'

A more marked contrast was observed in the supplementary Chinese newspaper corpus PDC2000, where 472 instances of *hou4guo3* (后果) were found. Of these 470 instances show a negative affective meaning, with the other two being neutral. The PDC2000 corpus shows two collocates with a minimum co-occurring frequency of 20, *yan2zhong4* 'serious, grave' and *zao4cheng2* (造成) 'cause'.

Like English *consequences*, all of the neutral occurrences of *hou4guo3* (后果) in LCMC were found in academic prose (e.g. 3a). *Hou4guo3* (后果) was also found to occur in contexts where an option between a desirable effect (e.g. 'peace') and an unpleasant consequence (e.g. 'disaster') is available, as shown in (3b). Note, however, that whilst *hou4guo3* (后果) can be used neutrally, the pattern meaning is quite

unambiguous – a negative evaluation – when it collocates with verbs like *zao4cheng2* (造成)/*dao3zhi4* (导致)/*zhi4shi3* (致使) 'cause' (see section 3.2).

(3)   a.   *shen1 ceng2ci4 ren4shi   de   hou4guo3   biao1zhi4-zhe*

                deep level   knowledge GEN consequence mark-ASP

                ge4ti3   ying4fu4 ying4ji1   de   neng2li1 (LCMC: J)

                individual cope-with emergence GEN ability

                'Deep level knowledge allows an individual to cope with emergences.'

        b.   *qi2   yin3qi3 de   hou4guo3   jiang1 bu4 shi4 he2ping2,*

                they cause   GEN consequence will   not be   peace

                er2 shi4 zai1nan4 (PDC2000)

                but be disaster

                'The consequences caused by any of such words and deeds will not be peace but a disaster.'

In contrast to the typically positive *jie2guo3* (结果) and the typically negative *hou4guo3* (后果), *shuo4guo3* (硕果) and *cheng2guo3* (成果) are inherently positive whereas *ku3guo3* (苦果) and *e4guo3* (恶果) are inherently negative, regardless of genre.[7] There are 4,572 instances of *cheng2guo3* (成果) and 109 instances of *shuo4guo3* (硕果) in the LCMC and PDC2000 corpora. The typical collocates of *cheng2guo3* (成果) include ***feng1shuo4*** 'rich and great', ***jiang3*** 'award', *zhuan3hua4* 'transform, turn into', *ke1ji4* 'science and technology', *yan2jiu4* 'research', *qu3de2* 'gain', ***you1xiu4*** 'excellent', ***gong4xian4*** 'contribution' and *sheng1chan3li4* 'productivity'. The significant collocates of *shuo4guo3* (硕果) include *lei3lei3* 'many times; countless' and *jie2chu1* 'yield'. *Cheng2guo3* (成果) is significantly more frequent than *shuo4guo3* (硕果), reflecting the fact that in the real world, results that

can be labelled as *shuo4guo3* (硕果) are considerably fewer than those labelled as *cheng2guo3* (成果). *Ku3guo3* (苦果) occurs 32 times and *e4guo3* (恶果) 42 times in the two Chinese corpora. All of these are negative, but no significant collocate was found for the two items.

Like the synonyms of *result* in English, the six near synonyms of *jie2guo3* (结果) in Chinese can be arranged on a semantic continuum, from positive to negative, as follows: *shuo4guo3* (硕果), *cheng2guo3* (成果), *jie2guo3* (结果), *hou4guo3* (后果), and *ku3guo3* (苦果)/*e4guo3* (恶果). In relation to English, it appears that Chinese is more sharply divided between the clearly negative and positive ends of the continuum. It is also important to note that unlike English, in which different forms of a lemma may have different collocates and semantic prosodies (e.g. *consequence* vs. *consequences* as noted earlier), Chinese does not have a rich morphology which can affect collocation and semantic prosody in this way.

3.2 The *cause* group

The negative semantic prosody of *CAUSE* has been widely observed (most notably by Stubbs 1995). In the FLOB and Frown corpora, there are 287 instances of *CAUSE* used as a verb. Of these, 223 occurrences have an unfavourable prosody, 56 are neutral while 8 cases are positive. The nominal collocates of *CAUSE* are quite revealing of what is typically associated with *CAUSE*: *damage*, *problems*, *pain*, *disease*, *distress*, *trouble*, blood, *concern*, *degradation*, events, *harm*, land, number, people, *pollution*, *suffering*, *anxiety*, condition, *death*, decision, *fear*, heart, human, *stress*, surprise, *symptoms*, time, water, women and years. The underlined items are all unpleasant. These examples of 'bad company' collocate with *CAUSE* so frequently that the central and typical use of this word shows a negative affective meaning.

When CAUSE is used neutrally, there is usually a contrast between something good and something bad, e.g. between 'pleasure' and 'pain' in (4). Also, considering the distribution of *consequences* across text categories, it is hardly surprising that the neutral use of CAUSE typically occurs in academic prose, which provides 28 out of 56 neutral occurrences, as shown in (5). Unfortunately, while it has been observed that collocation is associated with text types (e.g. Firth 1957: 195; Sinclair 1966: 429; Greenbaum 1974: 81; Partington 1998: 17; Hoey 2003), the link between semantic prosody and text type has largely been overlooked in previous research (though see Tribble 1999: 206 and Hoey 2003 for exceptions).

(4)   Utilitarianism asks the individual to aggregate the consequences of his or her actions for the promotion of pleasure and avoidance of pain, and demands that she or he should morally only follow that course which causes more pleasure than it does pain. (FLOB: J)

(5)   a.   These echoes are caused by reflection of the speaker's voice back from the distant receiving end […]. (FLOB: J)

        b.   A portion of the increase in phosphate was caused by an increase in phosphotyrosine. (Frown: J)

The apparently neutral or even positive occurrences of CAUSE are often offset by their longer co-texts, as shown in (6).

(6)   a.   His dress and the manner in which he carried himself caused her to think of a strutting peacock, a very vicious, very deadly one. (Frown: P)

        b.   That idea caused a brief, groundless burst of hope which brought her a little closer to real waking. (Frown: L)

In (6a), *think of a strutting peacock* appears quite neutral. Nevertheless, the neutrality is immediately traded off by *vicious* and *deadly*. Similarly, whilst *hope* in

(6b) is quite positive, the hope is turned into a fantasy or daydream when *groundless* is used.

The central and typical negative prosody, however, does not prevent CAUSE from co-occurring with words indicating positive results (e.g. *such pleasure, some easing of the confusion, some tumors to shrink or disappear*), although such cases are extremely rare (see the discussion of *zao4cheng3* (造成) 'cause'). Sometimes, a word with a strongly negative semantic prosody such as CAUSE is used to achieve an ironic effect.

(7)   Two million children could be dying of hunger in the Sudan, and that wouldn't cause a bump in consciousness. But Richard Johnson's wife seen with a long-haired stranger – now that was news! (Frown: K)

In (7), for example, the situation of two million children dying of hunger certainly should be the cause of 'a bump in consciousness'. In this sense, a 'bump in consciousness' is undoubtedly a desirable result. The conflict between this desirable result and the strongly negative word CAUSE triggers a sharply ironic effect: whilst trivial news like someone's wife having an affair spreads rapidly and is of general interest, truly important news such as two million starving children may pass unnoticed.

Whilst CAUSE and BRING about have a very similar denotational meaning, their collocational behaviour and semantic prosodies differ markedly. In contrast to the strongly negative prosody of CAUSE, the objects of BRING about more frequently refer to desirable, or at least non-negative situations (cf. Johns 1997). In the FLOB and Frown corpora, there are 37 occurrences of BRING about, of which 17 are positive and 14 are negative, with the remaining six being neutral. As the overall frequency is low in the two corpora, no significant collocate was found that indicated the nature of what was brought about. However, the pattern found in FLOB/Frown generally

conforms with the pattern in the BNC, where 2,150 instances of *BRING about* were found.[8] In the BNC, the most significant nominal collocates are *change* and *changes* (the MI scores for *changes* and *change* are 5.48 and 5.17 respectively, occurring 378 times as a collocate of *BRING about*), which can be either good or bad. It is also of interest to note that while *BRING about* and *CAUSE* are both significant collocates of *change(s)*, the former occurs significantly more frequently (436 times as a collocate, of which 378 times appearing as *BRING about*) than *CAUSE* (251 times) in the BNC. *Change(s)* collocating with *BRING about* typically shows a favourable evaluation whereas its combination with *CAUSE* typically shows an unfavourable evaluation. Other significant nominal collocates include **improvements**, *revolution*, **order**, *increase*, <u>*death*</u>, <u>*downfall*</u>, <u>*war*</u>, *government*, *situation*, *action*, **improvement**, *policy*, *reduction*, *result* and *state*. Of these, with the exception of the three underlined items which are obviously negative, the other collocates are either positive or neutral. Adjectival collocates are also positive (e.g. *new, major, significant*) or neutral (e.g. *social*, *economic*). It is clear from the frequency data in FLOB/Frown and significant collocates in the BNC that *BRING about* typically does not express a negative affective meaning. As such, those who support euthanasia prefer to use *BRING about death* over *CAUSE death* (see Louw 2000: 58).

In addition to *CAUSE* and *BRING about*, there are a number of phrasal verbs that share similar meanings but vary in semantic prosodies. They include e.g. *RESULT in/from*, *LEAD to*, *AROUSE* and *GIVE rise to*.[9] *RESULT in* and *RESULT from* are quite similar in affective meanings irrespective of some differences in frequency (the former is significantly more frequent than the latter) and syntax (event A results in event B whereas event B results from event A). On the semantic level, they are used less frequently than *CAUSE* to show a negative affective meaning and less frequently

than *BRING about* to show a positive evaluation. Their significant collocates include *change(s)*, as well as items of both positive (e.g. **growth** and **development**) and negative (e.g. *loss*, *injury*, *damage* and *death*) meanings. *LEAD to* is strikingly similar to *RESULT in/from* in affective meanings. The positive collocates include **understanding** and **development** while the negative collocates include *accident*, *loss*, *problems* and *disease*. Most collocates of *LEAD to* are neutral (e.g. *competition*, *increase*, *action*, *activity*, *conditions*, *courses* and *decision*). The neutral occurrences of *RESULT in/from* (33 out of 47) and *LEAD to* (45 out of 85) were typically found in academic prose. *AROUSE* and *GIVE rise to* are relatively infrequent. In FLOB/Frown, there are 29 instances of *GIVE rise to* (45% negative, 52% neutral and 3% positive; no significant collocate) and 17 occurrences of *AROUSE* (76% negative and 24% neutral, the only collocate being *suspicion*).

As such low frequencies may produce misleading results, we used four corpora of the Brown family (LOB, FLOB, Brown and Frown) to examine the distribution of *GIVE rise to* and *AROUSE* across meaning categories. As can be seen from Table 3, the increase in frequencies makes the distribution of the two phrasal verbs more balanced across the three meaning categories, but it does not lead to a drastic change in ordering. In these four corpora, no significant collocate was found which indicates the result referred to by *GIVE rise to*; only one such collocate, *suspicion*, was found for *AROUSE*.

[TABLE 3 NEAR HERE]

[TABLE 4 NEAR HERE]

The *CAUSE*-words in Chinese include *chan3sheng1* (产生, 361 instances in the LCMC corpus), *xing2cheng2* (形成, 334), *zao4cheng2* (造成, 208), *yin3qi3* (引起, 192), *dai4lai2* (带来, 131), *dao3zhi4* (导致, 79), *cu4shi3* (促使, 44), *zhi4shi3* (致使,

21

23), *yin3fa1* (引发, 11), *cu4cheng2* (促成, 11) and *niang4cheng2* (酿成, 4). These near synonyms differ in semantic prosodies. As with the CAUSE-words in English, we use supplementary data from PDC2000 when the overall frequency is very low (below 30). Table 4 shows the distribution of these near synonyms across meaning categories. Note that the frequencies of *zhi4shi3* (致使), *yin3fa1* (引发), *cu4cheng2* (促成) and *niang4cheng2* (酿成) as shown in the table were extracted from the PDC2000 corpus. It can be seen from the table that *zhi4shi3* (致使), *niang4cheng2* (酿成) and *zao4cheng2* (造成) are overwhelmingly negative; *yin3fa1* (引发) and *dao3zhi4* (导致) are very frequently used with a negative semantic prosody. In contrast, *cu4cheng2* (促成) and *cu4shi3* (促使) typically show a positive affective meaning. The collocates indicating what was referred to by these near synonyms are listed below (in descending order of frequency of co-occurrences).

- *Zhi4shi3* (致使): (collocates from PDC2000) *si3wang2* 'death', *wen4ti2* 'problem', *sun3shi1* 'loss', *chong1tu1* 'conflict', *shi4gu4* 'accident', *shi4jian4* 'incident; event', *po4huai4* 'destruction', *fu4dan1* 'burden'

- *Niang4cheng2* (酿成): (collocates from PDC2000) *da4huo4* 'disaster', *bei1ju4* 'tragedy', *shi4gu4* 'accident', *hou4guo3* 'consequence', *wei1ji1* 'crisis', *huan4* 'trouble', *chong1tu1* 'conflict'

- *Zao4cheng1* (造成): *sun3shi1* 'loss', *hou4guo3* 'consequence', *kun4nan2* 'difficulty', *jie2guo3* 'result', *ji2bing4* 'disease', *zhan4zheng1* 'war', *lang4fei4* 'waste', *sun3shi1* 'damage'

- *Yin3fa1* (引发): (collocates from PDC2000) *wen4ti2* 'problem', *mao2dun4* 'conflict', *chong1tu1* 'conflict', *wei1ji1* 'crisis', *shi4gu4* 'accident', *bao4zha4*

'explosion', *tong1huo4peng2zhang4* '(currency) inflation', *shi4jian4* 'incident; event', *fan4zui4* 'crime', *jiu1fen1* 'dispute', *si1kao3* 'thinking'

- *Dao3zhi4* (导致): (collocates from PDC2000) *xia4die1* 'decline', *li2hun1* 'divorce', *shi1wu4* 'error; fault', *e4hua4* 'deterioration', *si3wang2* 'death', *hou4guo3* 'consequence', *shi1bai4* 'failure', *jin3zhang1* 'tension', *ju2shi4* 'situation', *liu2shi1* 'loss', *bao4li4* 'violence'

- *Dai4lai2* (带来): *kun4nan2* 'difficulty', *wen4ti2* 'problem', *sun3shi1* 'loss', *ying3xiang3* 'influence'

- *Yin3qi3* (引起): *guan1zhu4* 'interest; attention', *ji2bing4* 'disease'

- *Chan3sheng1* (产生): *ying3xiang3* 'influence', *zuo4yong4* 'effect', *xi4tong3* 'system', **fa1zhan3** 'development', *bian4hua4* 'change', *ying4ji1* 'stress', *zong1jiao4* 'religion', *xiao4ying4* 'effect', *xin1li3* 'psychology; feeling', *huai2yi2* 'suspicion', *cuo4wu4* 'mistake; error', *gan3ying4* 'response; reaction', *ying4hua4* 'hardening', *wu4cha1* 'error', *qing2xu4* 'feeling', *jie2guo3* 'result'

- *Xing2cheng2* (形成): *jie2gou4* 'structure', *xi4tong3* 'system', *gui1mo2* 'scale', *guan1xi4* 'relationship', *ge2ju2* 'pattern', *ren4shi* 'knowledge', *wen4ti2* '(research) question', *dui4zhao4* 'contrast', *ji1lei3* 'accumulation', *ju2mian4* 'situation', *xi2guan4* 'habit', *yin4xiang4* 'impression', *ying3xiang3* 'influence', *ji1zhi4* 'mechanism', *ti3xi4* 'system', *guan1nian4* 'idea'

- *Cu4shi3* (促使): (collocates from PDC2000) **fa1zhan3** 'development', *jing4zheng1* 'competition', **jia1qiang2** 'strengthening', *zhuan3hua4* 'transformation', **zeng1zhang3** 'growth', **ti2gao1** 'increase', **jie3jue2** 'solution', **kai1fa1** 'development', **he2zuo4** 'cooperation'

- *Cu4cheng2* (促成): (collocates from PDC2000) **he2ping2** 'peace'

These collocates and the proportions for the negative/positive meaning categories are in conformance with each other. Both the proportional data and the collocates indicate that *zhi4shi3* (致使) and *niang4cheng2* (酿成) are essentially negative. Typical examples of the two items are given in (8).

(8)   (a)   *ren2men ke3yi3 zhao3chu1 xu3duo1 zhi4shi3 ta1 duo4luo4 de*

    people   can   find   many   cause   he degenerate GEN

    yuan2yin1 (LCMC: C)

    reason

    'We can find many causes for his degeneration.'

    (b)   *ru2guo3 yu4shang4 zhong4da4 qing2kuang4, zhe4yang4 cao3shuai4*

    if   meet-with critical   situation,   this   sloppy

    xing2shi4 nan2mian3 niang4cheng2 da4huo4 (LCMC: J)

    act   unavoidably cause   great-disaster

    'In critical situations, taking hasty action like this would inevitably lead to a great disaster.'

Zao4cheng2 (造成) is also overwhelmingly negative, as exemplified in (9a). It shows such a strongly negative prosody that even an apparently neutral result may become negative, as in (9b). In political writing, however, such a strongly negative word can be equally positive, as in *zao4cheng2* (cause) *yi1* (one) *ge4* (CL) <…> *liang2hao3* (good) *de* (GEN) *she4hui4* (social) *huan2jing4* (environment) 'create a favourable social environment, in which <…>'. Our explanation for this is that while the more frequently neutral synonym *xing2cheng2* (形成) collocates perfectly well with such a positive result, *zao4cheng2* (造成) involves more human effort (*xing2cheng2* (形成)

refs to a result coming as a natural course) and is thus more powerful in achieving the desired effect.[10]

(9)   (a)   *ni3 bi4xu1 dui4 ni3 zao4cheng2 de   yan2zhong4 hou4guo3*

you must for you cause   GEN serious   consequence

fu4ze2 (LCMC: K)

responsible

'You must be responsible for the serious consequences you have caused.'

(b)   *wo4shi4 de   chuang1hu4 mei2you3 guan1, bo2bo2 de*

bedroom GEN window   not   close, thin   GEN

chuang1lian2 zai4 ye4feng1 li3 piao1piao1fo2fo2, zao4cheng2

curtain   in   night-wind in flutter-flutter,   cause

yi1-zhong3 ji2 ju4 lang4man4 qing2diao4 de,   fei1dong4 de

one-CL   very have romantic   appeal   GEN, flying   GEN

yin4xiang4, zheng4 xiang4 nü3zhu3ren2 xi3nu4wu2chang2,

impression, just   like   hostess   changing-moods,

zao4dong4 bu4 ning2 de   xing4ge2   (LCMC: P)

restless   not   calm GEN disposition

'The window of the bedroom was open. The thin curtain was fluttering gently in the night wind, giving an impression of romantic appeal and flying, just like the restlessly changing moods of its hostess.'

The collocates and the proportions show that *dao3zhi4* (导致) and *yin3fa1* (引发) more often than not indicate an unfavourable evaluation and most of their collocates are negative while *dai4lai2* (带来), *yin3qi3* (引起) and *chan3sheng1* (产生) are less frequently used in a negative context. As such, *chan3sheng1* (产生) *cheng2guo3* (成

25

果) evaluates a result as positive, and *chan3sheng1* (产生) *hou4guo3* (后果) evaluates a result as negative while *chan3sheng1* (产生) *jie2guo3* (结果) can be good or bad. *Xing2cheng2* (形成) is typically used in a neutral context while *cu4cheng2* (促成) and *cu4shi3* (促使) most frequently show a positive semantic prosody, as reflected by the nature of their collocates.

The discussion of the *consequence* and CAUSE-groups of synonyms in English and their close translation equivalents in Chinese suggests that both languages exhibit features of semantic prosody, and that near synonyms are normally not collocationally interchangeable in either language as they show different semantic prosodies (consider CAUSE *a change* vs. BRING *about a change*, and *zao4cheng2* (造成) *jie2guo3* (结果) vs. *chan3sheng1* (产生) *jie2guo3* (结果)). The comparative analysis also shows that close translation equivalents display very similar collocational behaviour and semantic prosodies in both English and Chinese (consider the collocates of *consequences* vs. *hou4guo3* (后果), and CAUSE vs. *zhi4shi3* (致使)/*zao4cheng2* (造成)).

3.3. *Price* vs. *dai4jia4* (代价)

In this section, we will examine a group of synonyms related to *price* and *cost*. In English, they include, e.g. *at (the) price(s)/cost(s)/expense (of)*, *at (a) price/cost*, *pay a price/cost*, and *pay the price/cost of*. Recently, Barclays Bank in the UK used the slogan 'The personal loan with the personal price' to promote their personal financial services. The slogan, however, provoked a heated discussion among corpus linguists. Some of them asserted that *personal price* carries a strongly negative semantic

prosody and was inappropriate given the purpose of the advert. Others saw nothing wrong with this advertisement.[11]

It was suggested in the discussion that the potentially negative prosody of *personal price* in the advertisement might have been caused by a wider phraseology, i.e. its co-occurrence with PAY or *at*. Some linguists speculated that lexical choice would also affect semantic prosodies, e.g. *personal cost* would be more negative than *personal price*. Co-occurrences of PAY */at/with* with *price(s)/cost(s)* can denote either literal (i.e. in monetary terms) or metaphorical (i.e. related to a loss, damage or sacrifice) meanings. The former use is neutral whereas the latter shows a negative semantic prosody. Is it true, then, that the co-occurrence of PAY with *price(s)/cost(s)* is more likely to denote a neutral literal meaning than the co-occurrence of *at* with the two words? Or is it true that the co-occurrence of *cost(s)* with PAY/at/with is more likely to denote a negative metaphorical meaning than the co-occurrence of *price(s)* with PAY/at/with? In the four corpora of the Brown family, a total of 198 such instances were found, as shown in Table 5. Log-likelihood tests show that the differences in the frequencies of literal and metaphorical uses of PAY vs. *at*, and of *cost(s)* and *price(s)* are not statistically significant at all. This means that PAY and *at* on one hand, and *cost(s)* and *price(s)* on the other hand, are equally likely to denote a neutral literal meaning or a negative metaphorical meaning.

[TABLE 5 NEAR HERE]

[TABLE 6 NEAR HERE]

A breakdown of the combinations in Table 5 tells us more about their pattern meanings. In Table 6, the symbol @ represents a pre- or post-modifier, with or without a preceding article (excluding the pattern *at any/all/a price/cost*, which was counted separately). It can be seen from the table that patterns 6 – 9 are typically used

to denote a negative metaphorical meaning (e.g. 10), whereas patterns 1 – 5 can be used either literally or metaphorically. The word 'typically' is important here. Because of the low overall frequencies of patterns 5 – 9 in the four corpora, the literal use of these patterns is rare in our data. The BNC does provide a number of instances of the literal use. For example, 20 out of the 61 occurrences of *at any price* in the BNC denote a literal meaning. The patterns *at a price* and *at a cost* also typically denote a literal meaning when there is a post-modifier (e.g. *at a price specified as ten years' rental*, *at a price of $12.50 a share*, *at a cost of £8m*, *at a cost that is affordable by the taxpayer*). When co-occurring with *at*, *price(s)* and *cost(s)* are equally likely to convey a semantically neutral literal meaning (the difference is not statistically significant, with a log-likelihood score of 2.364 and a p value of 0.124), while in collocation with PAY, the literal use of *cost(s)* is significantly more frequent than *price(s)* (LL=5.869 and p=0.015). It is also interesting to note that when followed by *of*, pattern 2 typically selects the singular form *cost*, and *at a cost of* is most frequently followed by an amount of money (e.g. *at a cost of £ 1.5 million*) while *at the cost of* is normally followed by an adverse effect (e.g. *at the cost of his own life* and *at the cost of pollution*). When used metaphorically, the adjectival collocates of *price/cost* demonstrate a semantic preference for items indicating unpleasantness (e.g. *unpleasant*, *terrible* and *fearful*) or enormous size (e.g. *any*, *all*, *obvious*, *heavy*, *too great*, *too high*, *whatever* and *what a*). The prices and costs are also quite revealing of their nature, covering matters as diverse as *some efficiency*, *her honor*, *pollution*, *self-destruction*, *his own life* and *bringing endless disasters on ourselves*.

(10)   (a)   The politicians – all of them – are out for votes and power at any price.

(FLOB: F)

(b)   This the Air Ministry was determined to avoid at all costs as an inquiry would have revealed the true reasons behind the acquisition of the land. (FLOB: G)

(c)   He would bring the Somers safely into New York harbor but at a price. Dear God, at what a price. (Frown: P)

(d)   So he seems to have overcome dualism of Happiness and Duty but at a cost. (FLOB: D)

Co-occurrences of *with* with *price(s)* and *cost(s)* are infrequent in FLOB/Frown, where only three instances of *cost(s)* and no instance of *price(s)* were found.[12] Data from the BNC shows that the co-occurrences of *with* and *cost(s)* and *price(s)* are typically literal and neutral, as shown in (11). The BNC also shows that the significant collocates of the phraseology *with @ price* are typically neutral technical terms (e.g. *list/premium/suggested/guaranteed/exercise price*).

(11)   a.   […] Topaz 2 would be used in a program with an overall cost of $100 million. (Frown: B)

        b.   But be assured, I had no intention of trying to buy your favours with the price of a meal. (BNC)

The Chinese equivalent of *price/cost* when used in this sense is *dai4jia4* (代价), which can be used either neutrally to denote a literal meaning (i.e. in monetary terms) or negatively to express a metaphorical meaning. In modern Chinese, the use of *dai4jia4* (代价) is typically metaphorical and negative, because its literal use is replaced by a number of near synonyms, which are technical terms: *jia4ge2* (价格)/*jia4kuan3* (价款)/*jia4qian2* (价钱) 'price' and *cheng2ben3* (成本) 'cost'. Note, however, that in martial arts fiction, which is usually written in a form of vernacular Chinese, i.e. modern Chinese styled to appear like classical Chinese, *dai4jia4* (代价)

is typically neutral. In the LCMC corpus, five out of 32 instances of *dai4jia4* (代价) are neutral (e.g. 12), with four of them from category N (martial arts fiction) and one from category R (humour). Clearly, the largest proportion of the neutral use of *dai4jia4* (代价) in LCMC is attributable to the martial arts fiction category, where all of the four instances of *dai4jia4* (代价) were neutral.

(12)   a.   *wei4shen2me ni3   de   dai4jia4 zhi3you3*

              why   you GEN price   only-have

              wu3shi2-liang3? (LCMC: N)

              fifty-tael[13]

              'Why do you only charge 50 tael?'

    b.   *qi1zi: ni3 bu4bi4   dao4 yue4tai2 shang4 song4xing2 le,*

              wife: you need-not go   platform on   send-off   PRT,

              yao4 hua1 yi1-kuai4 qian2 mai3 yue4tai2piao4 de (LCMC: R)

              will cost one-dollar money buy   platform-ticket PRT

              'Wife: You don't have to go the platform to send me off, it costs one dollar to buy a platform ticket.'

              zhang4fu1: mei2you3-guan1xi4, zhi3 hua1 zhe4me xiao3 de   dai4jia4

              Husband: never-mind,   only cost so   small GEN price

              jiu4 neng2 ba3 ni3 song4-zou3, shi2zai4 tai4 zhi2de le

              then can   BA you send-off,   indeed   really worth PRT

              'Husband: "It's all right. It's worth such a small price to be able to send you off."'

In LCMC only two significant collocates were found for *dai4jia4* (代价), namely *fu4chu1* 'pay' and *tai4gao1* 'too big'. The essentially negative semantic prosody of *dai4jia4* (代价) observed in the LCMC corpus was confirmed by PDC2000, where

423 instances of *dai4jia4* (代价) were found. Of these 98% are used negatively, with only 10 instances being neutral (9 instances used as the titles of TV series and one instance in *dai4jia4quan4* 'voucher'). The modifying collocates extracted from PDC2000 include <u>*chen2zhong4*</u> 'heavy', *da* 'big', *yi1qie4* 'all', *gao1ang2* 'exorbitant', <u>*can3zhong4*</u> 'grievous', *ju4da4* 'enormous', *bu4xiao*3 'not small', *ang2gui4* 'expensive', *gao1* 'high', <u>*can3tong4*</u> 'painful', <u>*sheng1ming4*</u> 'life' and <u>*xue3*</u> 'blood'. It appears, then, that like its close equivalents (metaphorical use of *price/cost*), *dai4jia4* (代价) in Chinese is typically negative and its collocates exhibit a semantic preference for items showing unpleasantness and enormous size. One obvious exception to such semantic preference is (12b), where the husband uses it to achieve a humorous effect.

While our corpus data shows that *dai4jia4* (代价) is typically described as high, *Xian4dai4 Han4yu3 Ci2dian3* 'The Dictionary of Modern Chinese' (the 1983 edition), the most authoritative Chinese dictionary in China, gives a surprising example: *yong4* (use) *zui4xiao3* (smallest) *de* (GEN) *dai4jia4* (price) *ban4* (do) *geng4duo1* (more) *de* (GEN) *shi4qing2* (thing) 'to accomplish more at least cost'. There is nothing wrong with the example itself. Nevertheless, it does not reflect the central and typical use of *dai4jia4* (代价) in attested language data. As this is the only example the dictionary gives for *dai4jia4* (代价), the dictionary fails to reflect the typical use of the word and is therefore somewhat misleading.

In contrast with the negative *dai4jia4* (代价), all of the occurrences of *jia4ge2* (价格, 165), *jia4qian2* (价钱, 19), *jia4kuan3* (价款, 1) and *cheng2ben3* (成本, 129) in LCMC are used in a neutral literal sense. The neutrality of infrequent items (i.e. *jia4qian2* (价钱) and *jia4kuan3* (价款)) was also confirmed by PDC2000, where all

of the 111 instances of *jia4qian2* (价钱) and 46 instances of *jia4kuan3* (价款) were neutral. The collocates of these near synonyms include:

- *Jia4ge2* (价格): *di1* 'low', *shang4zhang3* 'increase', *pian2yi* 'cheap', *qi1huo4* 'futures', *mai4chu1* 'sell', *shi3yong4quan2* 'right of use', *xing4neng2* 'performance', *xia4die1* 'drop', *kou4chu2* 'deduct', *yuan2cai2liao4* 'raw material', *zi4ran2zi1yuan2* 'natural resource', *shang4sheng1* 'rise', *shang1pin3* 'commodity'

- *Jia4qian2* (价钱): (from PDC2000) **hao3** 'good', *mai4* 'sell', *gao1* 'high', *da4* 'big', *yuan2* (Chinese currency unit *yuan*)

- *Jia4kuan3* (价款): (from PDC2000) *suo3de2* 'earning', *he2tong2* 'contract', *gong1cheng2* 'project', *bian4mai4* 'sell off', *yuan2* (*yuan*)

- *Cheng2ben3* (成本): *jiang4di1* 'reduce', *cai2wu4* 'financial', *guan3li3* 'management'

Unlike the collocates of *dai4jia4* (代价), these items do not display preferences for items indicating unpleasantness and enormous size. It is clear that these near synonyms are collocationally different from *dai4jia4* (代价). Rather, they are used in the same way as the literal use of English *price/cost*.

4. Concluding discussion: some implications for language pedagogy

This paper has explored the collocational behaviour and semantic prosody of near synonyms from a cross-linguistic perspective. Our contrastive analysis shows that semantic prosody and semantic preference are as observable in Chinese as they are in English. As the semantic prosodies of near synonyms and the semantic preferences of their collocates are different, near synonyms are normally not interchangeable in either language. It can also be seen from the case studies that the semantic prosodies

observed in general domains may not apply to technical texts (consider, e.g. *consequences* and *hou4guo3* (后果) in academic prose). While English and Chinese are distinctly unrelated, the collocational behaviour and semantic prosodies of near synonyms are quite similar in the two languages (consider, e.g. CAUSE and *zao4cheng2* (造成)). This observation echoes the findings which have so far been reported for related language pairs, e.g. English vs. Portuguese (Sardinha 2000), English vs. Italian (Tognini-Bonelli 2001: 131-156), and English vs. German (Dodd 2000). While the corpus-based approach can only reveal but not explain such cross-linguistic similarity, at least part of the explanation, in our view, can be found in the common basis of natural language semantics – 'the conceptual system that emerges from everyday human experience' (Sweetser 1990: 1). However, as different languages can have different ranges of near synonyms (consider, e.g. *result/outcome* vs. *jie2guo3* (结果)/*cheng2guo3* (成果)/*shuo4guo3* (硕果), and *price(s)* vs. *jia4ge2* (价格)/*jia4qian2* (价钱)/*jia4kuan3* (价款)/*dai4jia4* (代价)), near synonyms and their close translation equivalents in different languages may also demonstrate, to some extent, different collocational behaviour and semantic prosody. A more general difference between English and Chinese is that collocation and semantic prosody may be affected by morphological variations in English but not in Chinese, which lacks such variation.

In addition to having implications for the contrastive study of language, our study also has at least three implications for language pedagogy. The first two relate to vocabulary teaching, particularly the teaching of near synonyms while the last relates to teaching language for specific purposes. We will discuss each in turn.

First, the importance of collocation and semantic prosody is well recognized in language learning (e.g. Partington 1998: 23-25; Hoey 2000, 2003; Hunston 2002: 142;

Altenberg and Granger 2001; Nesselhauf 2003). Yet the emphasis of this work has been monolingual. Our study shows that a contrastive analysis of collocation and semantic prosody would be useful to L2 learners. To illustrate this point, let us consider (13a), which was produced by a Chinese-speaking postgraduate of tourism. Because of the negative semantic prosody of *CAUSE*, Johns (1997) suggested revising the sentence as (13b):

(13) (a) Although economic improvement may be caused by tourism, the investment and operational costs of tourism must also be considered.

(b) Although tourism may lead to economic improvement, the investment and operational costs of tourism must also be considered.

In fact, inappropriate word choice arising from an ignorance of semantic prosody is not uncommon in Chinese learners of English, as shown in examples in (14), from the Chinese component (upper intermediate proficiency level) of the Longman Learners' Corpus.

(14) (a) The city caused him great interest, caused all citizens to grasp time and chances, to work for a better life.

(b) […] there are a lot of advantages are caused by them.

(c) During the past fifty years, the political, economic, and social changes in China have caused dramatic changes in people's lives.[14]

In Chinese, it is highly likely that the student who produced (13a) would express the idea in the first part of the sentence as *sui1ran2* (although) *lü3you2* (tourism) *ke3yi3* (may) *cu4shi3* (lead to; bring about) *jing1ji4* (economic) *hao3zhuan3* (improvement). In this context native speakers will naturally use *cu4shi3* (促使) rather than *zao4cheng2* (造成), because even though semantic prosody can only be reliably observed in a large number of key-word-in-centre (KWIC) concordances,

34

native speaker intuition certainly can detect the usage of a word at odds with its semantic prosody. Learners' L2 intuition, nevertheless, is inevitably less reliable than their L1 intuition.

If teachers compare the collocational behaviour and semantic prosody/preference of near synonyms in L1 and their close translation equivalents in L2, and make learners aware of L1-L2 differences, this should considerably reduce the number of errors from L1-L2 semantic prosody differences. While even distinctly unrelated languages such as English and Chinese may demonstrate similarity in the collocational behaviour and semantic prosody/preference of near synonyms, it is the teacher's responsibility to show learners which synonymous item in an L1 most closely matches which in an L2.

Second, as near synonyms usually differ in their collocational behaviour and semantic prosodies, the traditional practice of explaining meanings to learners by offering synonyms should be used with caution. Tognini-Bonelli (2001: 34) views using synonyms in this way to be a potential trap for learners because it emphasizes the denotational meaning of words rather than their usage. In practice, as can be seen in our case studies, near synonyms like *consequence* vs. *result* and *CAUSE* vs. *BRING about* typically operate in different contexts. We maintain that vocabulary, including synonymy, should be taught in a context that provides cues from which the learner can recall meaning and usage (cf. Harley 1996). As Weigand (1998) comments, to learn a language is to know how words are used and what utterances are used in specific situations. The learning tool best suited for this purpose is properly sorted KWIC (keyword-in-centre) concordances, as these allow the learner to observe repeated patterns and meanings, and thus help them to become aware of collocation and semantic prosody.

Third, learners must be made aware of the fact that collocation patterns and semantic prosodies can vary across text categories. The difference is more distinct between texts in general domains and technical or specialized texts. This observation is particularly relevant to the teaching and learning of domain-specific language – English for Specific Purpose (ESP), for example.

In conclusion, given the importance of contrastive analysis of collocation and semantic prosody to language learning and the present lack of this kind of research, our view is that there is a pressing need for the cross-linguistic study of collocation and semantic prosody to be pursued by researchers.

Acknowledgements

Notes

[1] In this paper Chinese characters are shown for synonyms under consideration. All other Chinese characters are converted into Romanized Pinyin symbols (with the numerals indicating tones).

[2] See http://www.ling.lancs.ac.uk/corplang/pdc2000/ for a description of the corpus.

[3] A word in small capitals refers to the lemma of the word, including its morphological variations. For example, CAUSE refers to *cause*, *causes*, *caused* and *causing*.

[4] In this list, only *accept* is a significant collocate as defined in this paper. In the BNC significant collocates (with a co-occurrence frequency of 20 or above) indicating actions include AVOID, ACCEPT, BE, CAUSE, CONSIDER, FACE, FOLLOW, HAVE, SUFFER.

[5] The sources of examples are given in the brackets following individual citations. In the grammatical glosses of Chinese examples, ASP stands for *aspect marker*, BA for the *ba*-construction with a preposed object, CL for *classifier*, GEN for *genitive*, and PRT for *particle*.

[6] Like *result, outcome* can also co-occur with *positive* and *negative*. In FLOB/Frown, one out of the three co-occurrences of *outcome(s)* with *positive* and one co-occurrence with *negative* are interpreted in a medical sense. One instance of *consequences* co-occurring with *negative*, but no co-occurrence with *positive* was found in FLOB/Frown. The BNC shows that *result(s)* and *outcome(s)* co-occur with *positive* (with a frequency of 161 and 41 respectively) more frequently than *negative* (74 and 11) whereas *consequence(s)* co-occurs with *negative* (36) more frequently than *positive* (10). The evidence from the BNC also reveals that *outcome(s)* and *consequence(s)*, when collocating with *positive* or *negative*, are not normally interpreted in a technical sense; when *result(s)* collocates with *negative*, it is more likely to refer to test results in a technical sense.

[7] As the inherent affective meanings (connotations) are not affected by collocates, we will not discuss these words in detail.

[8] To take into account instances such BRING *it/this about*, occurrences with BRING and *about* separated by one word were also included.

[9] The instances of LEAD *to* indicating physical locations (e.g. *the narrow hall that led to the kitchen*) are irrelevant here. Similarly, the occurrences of AROUSE *somebody* (from sleep or sexually, etc) are discarded.

[10] Greenbaum (1974: 89) observes that *entirely* collocates predominantly with verbs of agreeing (at least in British English) and disagreeing while the dominant feature common to most verbs collocating with *utterly* is their negative implication. Nevertheless, 'if we say *I utterly agree with you*, the usual negative implication makes what we say more emphatic than if we say *I entirely agree with you* (compare *terribly bright*)'.

[11] See Corpora Mail List Archive between October 17 and 22, 2003 for details.

[12] Occurrences such as *compared with an H-registration price of £12,900*, where *with* collocates with *compared*, were not counted here.

[13] In former times *tael* was a monetary unit for silver in weight.

[14] While *dramatic changes* in (14c) can be good or bad, all of the examples given showed that the changes were evaluated as desirable.

References

Altenberg, B. and Granger, S. 2001. 'The grammatical and lexical patterning of MAKE in native and non-native student writing'. *Applied Linguistics* 22/2: 173-195.

Collins. 2000. *Collins English Dictionary* (5th ed.). London: HarperCollins.

Conzett, J. 1997. 'Integrating collocation into a reading and writing course' in J. Coady and T. Huckin (eds.) *Second Language Vocabulary Acquisition*. Cambridge: Cambridge University Press.

Dodd, B. 2000. 'Introduction: the relevance of corpora in German studies' in B. Dodd (ed.) *Working with German Corpora*, pp. 1-39. Birmingham: University of Birmingham Press.

Firth, J. 1957. *Papers in Linguistics*. Oxford: Oxford University Press.

Firth, J. 1968. 'A synopsis of linguistic theory 1930–1955' in F. R. Palmer (ed.) *Selected Papers of J. R. Firth 1952–59*, pp. 1–32. Bloomington: Indiana University Press.

Francis, W., and Kučera, H. 1964. *Manual of Information to Accompany a Standard Sample of Present-day Edited American English, for Use with Digital Computers*. Providence: Brown University.

Greenbaum, S. 1974. 'Some verb-intensifier collocations in American and British English'. *American Speech* 49: 79-89.

Halliday, M. A. K. 1966. 'Lexis as a linguistic level' in C. Bazell, J. Catford, M. A. K. Halliday and R. Robins (eds.) *In Memory of J.R. Firth*, pp. 148-162. London: Longman.

Halliday, M. A. K. 1976. *System and Function in Language*. Oxford: Oxford University Press.

Harley, B. 1996. 'Teaching vocabulary: an exploratory study of direct techniques'. *Canadian Modern Language Review* 53: 281-304.

Hayakawa, S. 1994. *Cassell Guide to Related Words*. London: Cassell.

Hoey, M. 1991. *Pattern of Lexis in Text*. Oxford: Oxford University Press.

Hoey, M. 2000. 'A world beyond collocation: New perspectives on vocabulary teaching' in M. Lewis (ed.) *Teaching Collocations*, pp. 224-245. Hove: Language Teaching Publications.

Hoey, M. 2003. 'Lexical priming and the properties of text'. URL: www.monabaker.com/tsresources/LexicalPrimingandthePropertiesofText.htm. Accessed on January 9, 2004.

Hundt, M., Sand, A. and Siemund, R. 1998. *Manual of Information to Accompany the Freiburg-LOB Corpus of British English*. Freiburg: University of Freiburg.

Hundt, M., Sand, A. and Skandera, P. 1999. *Manual of Information to Accompany the Freiburg-Brown Corpus of American English*. Freiburg: University of Freiburg.

Hunston, S. 2002. *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.

Johansson, S., Leech, G. and Goodluck, H. 1978. *Manual of Information to Accompany the Lancaster-Oslo/Bergen Corpus of British English, for Use with Digital Computers*. Oslo: University of Oslo.

Johns, T. 1997. 'Cause v. lead to v. bring about'. URL: www.eisu.bham.ac.uk/Webmaterials/kibbitzers/kibbitzer24.htm. Accessed on January 9, 2004.

Krishnamurthy, R. 2000. 'Collocation: from *silly ass* to lexical sets' in C. Heffer, H. Sauntson and G. Fox (eds.) *Words in Context: A Tribute to John Sinclair on his Retirement*. Birmingham: University of Birmingham.

Longman. 1999. *Longman Dictionary of English Language and Culture*. Harlow: Longman.

Louw, B. 1993. 'Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies' in M. Baker, G. Francis and E. Tognini-Bonelli (eds.) *Text and Technology: In Honour of John Sinclair*, pp. 157-76. Amsterdam: John Benjamins.

Louw, B. 2000. 'Contextual prosodic theory: bringing semantic prosodies to life' in C. Heffer, H. Sauntson and G. Fox (eds.) *Words in Context: A Tribute to John Sinclair on his Retirement*. Birmingham: University of Birmingham.

Mackin, R. 1978. 'On collocations: "Words shall be known by the company they keep"' in P. Strevens (ed.) *In Honour of A. S. Hornby*, pp. 149-165. Oxford: Oxford University Press.

McEnery, A. and Wilson, A. 2001. *Corpus Linguistics* (2nd ed.). Edinburgh: Edinburgh University Press.

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: 361-378.

Merriam-Webster. 1991. *The Merriam-Webster Thesaurus*. Springfield: Merriam-Webster Inc.

Nesselhauf, N. 2003. 'The use of collocations by advanced learners of English and some implications for teaching.' *Applied Linguistics* 24/2: 223-242.

Partington, A. 1998. *Patterns and Meanings*. Amsterdam: John Benjamins.

Partington, A. 2004. '"Utterly content in each other's company": Semantic prosody and semantic preference'. *International Journal of Corpus Linguistics* 9/1: 131-156.

Sardinha, T. 2000. 'Semantic prosodies in English and Portuguese: A contrastive study'. *Cuadernos de Filología Inglesa* 9/1: 93-110.

Schmitt, N. and Carter, R. 2004. 'Formulaic sequences in action: An introduction' in N. Schmitt (ed.) *Formulaic Sequences*, pp. 1-22. Amsterdam: John Benjamins.

Sinclair, J. 1966. 'Beginning the study of lexis' in C. Bazell, J. Catford, M. A. K. Halliday and R. Robins (eds.) *In Memory of J. R. Firth*, pp. 148-162. London: Longman.

Sinclair, J. 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

Sinclair, J. 1996. 'The search for units of meaning'. *Textus* IX: 75–106.

Sinclair, J. 1998. 'The lexical item' in E. Weigand (ed.) *Contrastive Lexical Semantics*, pp.1–24. Amsterdam: John Benjamins.

Stubbs, M. 1995. 'Collocations and semantic profiles: on the cause of the trouble with quantitative methods'. *Function of Language* 2/1: 1-33.

Stubbs, M. 1996. *Text and Corpus Linguistics*. Oxford: Blackwell.

Stubbs, M. 2001a. 'On inference theories and code theories: Corpus evidence for semantic schemas'. *Text* 21/3: 437-465.

Stubbs, M. 2001b. *Words and Phrases*. Oxford: Blackwell.

Stubbs, M. 2002. 'Two quantitative methods of studying phraseology in English'. *International Journal of Corpus Linguistics* 7/2: 215-244.

Sweetser, E. 1990. *From Etymology to Pragmatics*. Cambridge: Cambridge University Press.

The Chinese Academy of Social Science. 1983. *Xian4dai4 Han4yu3 Ci2dian3* (*Dictionary of Modern Chinese*). Beijing: Commercial Printing House.

Tognini-Bonelli, E. 2000. 'Corpus classroom currency'. *Naujoji Metodologija* 24: 205-244.

Tognini-Bonelli, E. 2001. *Corpus Linguistics at Work*. Amsterdam: John Benjamins.

Tribble, C. 1999. *Writing Difficult Texts*. PhD thesis. Lancaster University.

Urdang, L. 1986. *Longman Synonym Dictionary*. Harlow: Longman.

Urdang, L. 1991. *The Oxford Thesaurus*. Oxford: Clarendon Press.

Weigand, E. (ed.) 1998. *Contrastive Lexical Semantics*. Amsterdam: John Benjamins.

Table 1 Examples of semantic prosodies

Author	Negative prosody	Positive prosody
Sinclair (1991)	BREAK out HAPPEN SET in	
Louw (1993, 2000)	bent on build up of END up *verb*ing GET oneself *verb*ed a recipe for	BUILD up a
Stubbs (1995, 1996, 2001a, 2001b)	ACCOST CAUSE FAN the flame signs of underage teenager(s)	PROVIDE career
Partington (1998)	COMMIT PEDDLE/peddler dealings	
Hunston (2002)	SIT through	
Schmitt and Carter (2004)	bordering on	

Table 2 Distribution of *consequence* across meaning categories in FLOB/Frown

Pattern	Negative	Neutral	Positive
as a consequence	6	7	4
in consequence (of)	8	3	1
consequence	27	7	6
consequences	85	20	1
consequent(ly)	15	73	5

Table 3 Distribution of *CAUSE*-words across meaning categories in English

Synonyms	Negative	Positive	Neutral
CAUSE	223 (78%)	8 (3%)	56 (19%)
AROUSE	45 (65%)	10 (15%)	14 (20%)
LEAD to	141 (49%)	65 (22%)	85 (29%)
RESULT in/from	84 (47%)	49 (27%)	47 (26%)
GIVE rise to	27 (46%)	8 (13%)	24 (41%)
BRING about	14 (38%)	17 (46%)	6 (16%)

Note: Frequencies for *GIVE rise to* and *AROUSE* include occurrences in FLOB, Frown, LOB and Brown.

Table 4 Distribution of *CAUSE*-words across meaning categories in Chinese

Synonyms	Negative	Positive	Neutral
zhi4shi3 (致使)	569 (99%)	0	3 (1%)
niang4cheng2 (酿成)	92 (98%)	2 (2%)	0
zao4cheng2 (造成)	190 (91%)	3 (2%)	15 (7%)
yin3fa1 (引发)	523 (77%)	31 (5%)	123 (18%)
dao3zhi4 (导致)	60 (76%)	2 (3%)	17 (21%)
dai4lai2 (带来)	64 (49%)	36 (27%)	31 (24%)
yin3qi3 (引起)	83 (43%)	28 (15%)	81 (42%)
chan3sheng1 (产生)	111 (31%)	88 (24%)	162 (45%)
xing2cheng2 (形成)	34 (10%)	85 (26%)	215 (64%)
cu4shi3 (促使)	2 (5%)	26 (59%)	16 (36%)
cu4cheng2 (促成)	2 (1%)	171 (98%)	2 (1%)

Table 5 *PAY/at* and *price(s)/cost(s)* in FLOB/Frown

Word	Type of use		LL score	Sig. level
	Neutral literal meaning	Negative metaphorical meaning		
PAY + price(s)/cost(s)	25	14	0.000	1.000
at + price(s)/cost(s)	100	56		
PAY/at/with + cost(s)	67	41	0.459	0.498
PAY/at/with + price(s)	60	30		

Table 6 Typical patterns of *price/cost*

No.	Pattern	Neutral literal meaning		Negative metaphorical meaning	
		Frequency	Percentage	Frequency	Percentage
1	at @ price(s) (of)	46	84%	9	16%
2	at @ cost(s) (of)	52	72%	20	28%
3	PAY @ price(s) (of)	12	50%	12	50%
4	PAY @ cost(s) (of)	13	87%	2	13%
5	with @ cost(s)	2	67%	1	33%
6	at any price	2	25%	6	75%
7	at any cost/at all costs	0	0%	16	100%
8	at a price	0	0%	3	100%
9	at a cost	0	0%	2	100%