

# Corpus-Based Study of Two Synonyms—Obtain and Gain

## Bing-jie GU Ningbo Dahongying University, Ningbo, China

With the advancement and prevalence of computer engineering, corpus-based approach of language analysis is on the rise. In this study, corpus-based approach is being adopted to compare two verbal synonyms—obtain and gain in terms of genre, colligation, collocation, and semantic prosody. The online tools, such as Sketch Engine, BNC Web, and Just the Word, are been adopted. The differences of two synonyms are as following: Noun is more collocated with obtain and passive voice pattern with preposition is more widely used. The verb form gain is collocated with abstract noun and most of them are endowed with the positive semantic prosody. It is also found that Oxford Dictionary failed to add the semantic prosody and dropped the frequently used collocation such as "obtained by pretence". For ESL (English as Second Language) teachers, they should be cautious in the traditional practice of explaining meaning to leaners by offering synonyms. Moreover, semantic prosody should be taken into account in translating English into Chinese.

Keywords: synonyms, colligation, collocation, semantic prosody

#### Introduction

With the advancement and prevalence of computer engineering, corpus-based approach of language analysis is on the rise. Comparing with the traditional approach, corpus-based approach can provide the linguists with improved reliability because it attaches great importance to empirical data and assists researchers to find differences that intuition alone may not perceive (Francis, Hunston, & Manning, 1996).

The aim of the essay is to adopt the corpus-based approach to discover the difference of two synonyms words (*gain* and *obtain*) in terms of genre, colligation, collocation, and semantic prosody. It can be divided into three parts. The first part is the brief literature review on synonyms study and the theory on word meaning exploration under corpus-based approach. The second part is the detailed study of the two synonyms words *gain* and *obtain* using online corpora tools *Sketch Engine*, *BNC Web*, and *Just the Word*. The rational and process will be further discussed. The third part is conclusion and possible implication of this study to the fields, such as dictionary writing, English language education and translation.

#### **Literature Review**

In the online Oxford Dictionary (2005), synonyms are defined as "a word or phrase that means exactly or nearly the same as another word or phrase in the same language". But these equivalencies can be very misleading, because synonymous words are typically used in different ways and convey different connotations. Corpus-based analyses are particularly well suited to unveil systematic differences, ranging from register difference to association with other collocations (Biber, Conrad, & Reppen, 2006, p. 43).

Bing-jie GU, lecturer, master, College of Humanity, Ningbo Dahongying University, Ningbo, China.

To be specific, corpora, known as a body of naturally occurring language, provides authentic texts which are sampled to be representative of a particular language (McEnery, XIAO, & Tono, 2006, p. 5). What is more, corpus-based analysis makes extensive use of computers to conduct quantitative tests on linguistic features, such as frequency and mutual information, which will be discussed in the second part. In addition to its linguistic feature, the non-linguistic features, such as varieties defined by register and periods of time, can be also found in corpus (Biber, Conrad, & Reppen, 2006, p. 7). One frequently cited limitation of corpus is that it cannot give information or explanations, though it can provide evidence for hypothesis (Hunston, 2002, p. 23). To address this problem, functional interpretations can be used to complement the quantitative analysis and to explain the statistics.

In terms of meaning comparison, Sinclair (2004) proposed that a word, as the unit of meaning, is related with other words around it (p. 27). In corpus, we can find out the meaning from the discourse it extracted from (Halliday, Teubert, & Cermakova, 2005, p. 105). According to Sinclair (1996), the internal structure of word has four parameters, which take different values and go from concrete to abstract: colligation, collocation, semantic preference, and semantic prosody. First, colligation is the concept proposed by Firth (1957) to refer to "the interrelation of grammatical categories in syntactical structure" (p. 12). Second, collocation, which defined as the items in the environment set by the span (McEnery & Hardie, 2012, p. 107), is widely used in synonymy comparison. In detailed analysis, there are two major approaches: "collocation-via-significance" as oppose to "collocation-via-concordance" (McEnery & Hardie, 2012, pp. 126-127). The former one depends on more rigorous inferential statistical tests than simple frequency counts and is now extensively used in collocation analysis (XIAO, 2008). Third, in terms of semantic preference, Stubbs (2001) defined it as the relation, not between individual words, but between a lemma or word form and a set of semantically related words. Fourth, words and phrases are said to have a negative or positive semantic prosody if they typically co-occur with units that have a negative or positive meaning according to Stubbs (1996, p. 176). If both positive and negative collocates exist in the context, the word can be said to bear a neutral or mixed semantic prosody. Notably, semantic prosody is a concept rooted in the concordance-based analysis of collocation (McEnery & Hardie, 2012, p. 136).

## The Corpus-Based Analysis of Two Synonyms

In this part, two synonyms *gain* and *obtain* will be studied in adoption of corpus-based analysis. *Obtain* is the verb while *gain* can be used as verb and noun. The focus is on the verb usage to facilitate the comparison. Based on the Oxford Dictionary, *obtain* and *gain* both ranked in top 1000 frequently used words. *Obtain* means "get and acquire something (wanted or desirable)" and *gain* means "obtain and secure" with additional meaning of *increase*, typically followed by *weight* or *speed*. The differences in genre distribution, colligation, collocation, and semantic prosody will be analyzed based on the three online corpora tools: *Sketch Engine*, *BYU-BNC*, and *Just the Word*. BNC (British National Corpus) will be used since its gold *Oxford Dictionary* standard among corpora of British English gives the large size, level of annotation, and availability (Anderson & Corbett, 2009, p. 10). BNC Web offers standard query of concordance, including user-friendly interface of collocation. In *Sketch Engine*, BNC is also one of the sub-corpora and *Sketch Diff* function offers collocation difference in a straightforward setting. *Just the Word* offers the colligation with frequency and it is also based on BNC. The detailed analysis with rationales is as followings.

#### **Genre Difference**

In this study, *Sketch Engine* is adopted to exam the specific domain in text types. Figure 1 and Figure 2 demonstrate the genre comparison of *obtain* and *gain* in terms of raw frequency and relative text type frequency in BNC. Both "*obtain*" and "*gain*" are typed in lemma bar as verb. Frequency limit of 5 is chosen. *Rel* (%) (the number of relative text type frequency) means relative frequency of the query result divided by relative size of the particular text type. The figure manifests that *obtain* is 3.4 times as common in natural and pure science of written English and 2.1 times as common in applied science than in the whole corpus. But, it is less frequently used in world affairs. *Gain* is more frequently used in commerce and finance, world affairs, and social science than in the whole corpus. In addition, *gain* is twice more than *obtain* in belief and thought. The major reason why these two synonyms have stark contrast in text domain distribution lies in the collocation part.

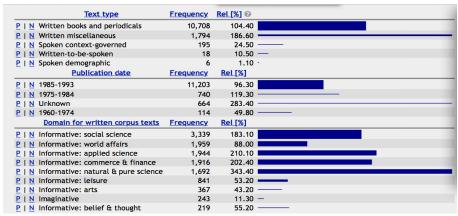


Figure 1. The genre of Obtain.

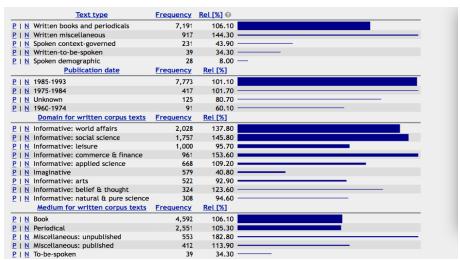


Figure 2. The genre of Gain.

## **Colligation Difference**

In terms of colligation, *obtain* and *gain* as verbs are categorized in the following patterns retrieved from online colligation website, *Just the Word* (see Figure 3):

Colligation	Frequency of Obtain	Frequency Ratio of Obtain	Frequency of <i>Gain</i>	Frequency Ration of Gain
V + n	9697	66%	6591	65%
V + prep + object	2479	17%	751	7.4%
n + V	1345	9%	1771	17.4%
adv + V	944	6.4%	979	9.6%
others	258	1.8%	87	0.9%
Total	14723	100%	10179	100%

Figure 3. The colligation difference of two words.

It can be found that *obtain* and *gain* mainly are collocated with object noun, noun subject, and adverb. In terms of similarity, these two words share the similar frequency ratio in the pattern of "v+obj n" and "v+adv". However, the difference in colligation lies in "subj+v" and "v+prep+object". It shows that *obtain* is more frequently used in the passive voice. The reasons will be also analyzed in the collocation part.

## **Collocation Difference**

The first part will focus on the pattern of "v+n" and "v+prep+obj" to see the collocation with noun given the high frequency. To put into practice, *gain* and *obtain* as verbs shall be typed {gain/V} and {obtain/V} as lemmas on BNC Web in a 0-5 span, though the drawback of including some unnecessary words shall be aware. Tag of noun will be chosen. When it comes to the significance of collocation, mutual information (MI) and *t* test are used together to measure collocation strength in consideration of corpus size (McEnery, XIAO, & Tono, 2006, p. 56). Hunston (2002, p. 71) proposed an MI score of 3 or higher to be taken as evidence that two items are collocated. A *t* score of 2 or higher is normally considered to be statistically significant. In this study, MI3 and *t* score higher than 2 are both examed to find the noun collocation of both words. Figure 4 shows the noun collocation difference of these two verbs.

Ob	tain:						Gair	n:					
No.	Word₽	Total	Expected	Observed	In	MI3	No.	Word₽	Total	Expected	Observed	In	MI3 +
		No. in	Collocate	Collocate	No.	Value ₽			No. in	Collocate	Collocate	No.	Value ₽
		Whole	Frequency	Frequency	of				Whole	Frequency	Frequency	of	
		BNC₽			texts				BNC₽			texts	
1⇔	Deception @	668₽	0.350₽	152₽	26₽	23.259₽	1₽	access₽	9,867₽	3.571 ₽	349₽	242₽	23.505 +
2₽	Information	38,266	20.041	370₽	221₽	21.269₽	2€	Experience @	18,593∉	6.729₽	281₽	219₽	21.653+
3₽	pretences₽	96₽	0.05₽	37₽	6₽	19.942₽	3₽	Insight ₽	1,404	0.508₽	94₽	81₽	20.641
4₽	Permission 4	3,164	1.657₽	110₽	97₽	19.615₽	4.	Confidence &	6,853₽	2.484₽	153₽	130₽	20.460₽
5₽	Consent ₽	3,233₽	1.693₽	87₽	52₽	18.569₽	5₽	momentum₽	952₽	0.345₽	67₽	60₽	19.736₽
6₽	Approval ₽	3,873₽	2.028₽	77₽	57₽	17.78₽	6₽	Reputation @	3,653₽	1.322₽	102₽	85₽	19.615₽
7₽	Licence &	3,551₽	1.860₽	71₽	53₽	17.554₽	7₽	control₽	23,539	8.519₽	188₽	144₽	19.573₽
8₽	copy₽	4,903₽	2.568₽	79₽	68₽	17.551₽	. 8€	support₽	20,948	7.581₽	179₽	140₽	19.529₽
9₽	possession@	3,122₽	1.635₽	58₽	20₽	16.8646	. 9₽	foothold₽	170₽	0.062₽	33₽	30₽	19.156₽
10₽	certificate₽	2,852₽	1.494₽	55₽	37₽	16.765₽	10₽	understanding	7,447₽	2,695₽	115₽	87₽	19.106
11₽	Injunction @	801₽	0.420₽	35₽	25₽	16.641₽	11₽	acceptance₽	2,629₽	0.951₽	81₽	77₽	19.091₽
12₽	Evidence 🕫	21,137	11.070₽	99₽	52₽	16.420₽	12₽	entry₽	5,082₽	1.839₽	92₽	77₽	18.692
13₽	advice₽	10,303	5.396₽	75₽	49₽	16.255₽	13₽	Recognition @	5,736₽	2.076₽	90₽	82₽	18.412₽
14₽	Property ₽	12,486	6.539₽	77₽	27₽	16.091₽	14₽	advantage₽	7,199₽	2.605₽	97₽	84₽	18.418
15₽	data₽	18,037	9.447₽	87₽	66₽	16.089₽	15₽	independence₽	4,415₽	1.598₽	81₽	45₽	18.343₽
16₽	results₽	14,638	7.666₽	76₽	63₽	15.805₽	16₽	ground₽	14,982	5.422₽	118₽	103₽	18.209₽
17₽	order₽	32,338	16.937₽	97₽	73₽	15.718₽	17₽	popularity₽	1,328₽	0.481₽	48₽	44₽	17.812
18₽	degree₽	9,814	5,140₽	61₽	45₽	15.431₽	18₽	seats₽	4,594₽	1.663₽	68₽	47₽	17.529₽

Figure 4. The noun collocation difference of two words.

Despite of the apparent similarity in meaning, the typical collocation of each verb differs to quite a considerable degree. From the chart, it can be found that *obtain* is more collocated with something taking in material form, especially the ones showing acknowledgment, while *gain* is more contingent with abstract concept, which also takes time and effort to get. It reveals that *obtain* is more used in the physical sense while *gain* is more used in metaphoric sense.

To be specific, in the list of *obtain*, words like *information*, *property*, *copy*, and *possessions* are concrete concept. The nouns *approval*, *consent*, and *permission* all show acknowledgement while *license*, *certificate*, and *degree* are in material forms and show acknowledgement. It is surprising to find that *deception* and *pretences*, which are abstract concept with negative denotation, ranked first and third on collocation list. Observing the concordance lines in KWIC (key word in context, see Figure 5), we can easily find the following two patterns: "obtain money/property by deception" and "obtain by deception". It is a criminal charge and is quite frequently used in academic law English (48 frequency of 152). But it should be pointed out that they are not adopted as an example in Oxford Dictionary. Similar pattern can be found in collocation of *pretences*: "obtain goods/property/money by false pretences". Among frequency of 37, 34 of them are used in academic law as a criminal charge as well. It also explains in some extent the pattern of "v+prep+object" is more frequently used in *obtain* as indicated in chart.



Figure 5. Concordance lines of "obtain by deception".

From the evidence of those collocations of gain, nouns like confidence, insight, reputation, and momentum are abstract concept with positive denotation, which also contains the meaning of "take time and effort to get". It can be inferred that the typical meaning of gain is metaphoric, rather than literal. At the same time, it is surprising to find that the nouns with concrete meanings, such as seat and ground, are also on the list, though they ranked in the end. The abstract meaning underpinned seat and ground can be found out after exploring the subject in the concordance lines (see Figure 6). For example, "gain seats" is predominantly used in politics with party name as the subject and exact number of seat following the verb gain. It refers to the power of the party and in alignment with the other abstract nouns in some degree. Statistically speaking, only one concordance line in 68 is used in business setting. The collocations of gain ground also share similar usage (see Figure 7). For example, the subjects of attitude, opinion, and impression in the following concordance lines are abstract concepts. In addition, the subjects of 23 concordances among 118 are related with one's thought, such as idea, view, and opinion. It justifies the more frequent distribution of gain in the domain of belief and thought in certain degree. In this sense, the meaning of ground goes beyond the physical level to correspond with the meaning of the subjects.

Labour needs to gain 97 seats and Peterborough is 97th on its target list.

THE Conservatives confounded all predictions in Scotland, gaining two seats and holding Stirling against a fierce Labour challenge. In the West Midlands, Labour gained six seats on a swing of 3.1 per cent but needed three more for victory.

PERSONNEL Personnel chiefs are gaining seats on company boards

Figure 6. Concordance lines of "gain seats".

In the age of AIDS and other incurable forms of venereal disease, new moral attitudes are gaining ground, and juries might well be led towards the view that explicit sexual depictions are degrading to such an extent as to exercise a depraving and corrupt influence on future conduct.

But then, as Blackstone pointed out, the opinion gained ground that it would be to the advantage of trade, and of creditors in general, if debtors outside the scope of the bankruptcy laws who were on the point of insolvency could also surrender their property for the benefit of their creditors, and in return be protected from legal process.

An impression has undoubtedly gained ground — which I do not think is fair to the Wolfenden Committee — that the Committee desired to legalise homosexual conduct.

This idea, too, is gaining ground across Europe.

The idea that older people should leave the labour force at some point gained ground in the earlier part of the twentieth century with the popularity of ideas about scientific management, which implied that older people were bound to be inefficient workers.

In private, however, the view is gaining ground that eventual membership of the Community by those East European countries that want to join it would be a good idea

Figure 7. Concordance lines of "gain ground".

The *Word Sketch Diff* in *Sketch Engine* website echoes the findings above. In the searching bar, lemma verbs and BNC are chosen and following chart (see Figure 8) shows the difference in terms of noun objects. It means that it is more usual to say "obtain data" and "obtain copy" than "gain data" or "gain copy", while it is more fluent to say "gain confidence" and "gain reputation". It to some extent explains the higher frequency of *obtain* used in natural and pure science since it is more collocated with words showing data and results.

Collocation(Noun)	Frequency	Frequency	Collocation	Collocation
	of	of	Significance	Significance
	Obtain	Gain	of	of Gain (MI)
			Obtain(MI)	
data	141	0	8.4	
copy	98	0	8.1	
license	72	0	7.9	
Result	277	6	9.1	3.7
Consent	129	6	8.8	4.8
Information	411	63	9.3	6.8
Permission	109	16	8.4	6.1
access	39	341	6.6	10.1
Insight	13	115	5.5	9.1
Experience	13	262	4.9	9.5
Momentum	0	62		8.3
Independence	0	72		8.4
Acceptance	0	68		8.4
Ground	0	111		8.5
Reputation	0	85		8.5
Confidence	0	113		8.8

Figure 8. Comparison of noun collocation.

The second part of difference lies in collocation with subjects. The *Stretch Engine* shows the difference in following chart (see Figure 9). It can be observed that the subject of *obtain* can be methods, such as *fraud*, *deception*, and *means*, or people in law or business field, such as *purchaser*, *buyer*, *plaintiff*, and *the accused*. It once again demonstrates the relevance of *obtain* with the domain of law and business. The subjects of *gain* are abstract concepts, such as *experience*, *knowledge*, *impression*, and *insight*, or related with politics and finance, such as *democrats*, *share*, and *benefit*. It in some degree explains the more frequent distribution of *gain* in world affairs and economics and finance.

Collocation(Subject)	Frequency	Frequency	Collocation	Collocation
	of	of	Significance	Significance
	Obtain	Gain	of	of Gain (MI)
			Obtain(MI)	
fraud	21	0	8.6	
Deception	16	0	8.3	
mean	20	0	7.8	
Accused	12	0	7.7	
Plaintiff	18	0	7.6	
Purchaser	7	0	6.8	
Buyer	9	0	6.7	
Result	26	0	6.7	
Sample	8	0	6.4	
Expansion	5	0	6.4	
Confidence	0	7		6.4
Skill	0	9		6.5
democrats	0	8		6.6
Share	0	17		6.5
benefit	0	14		7.0
Expertise	0	10		7.1
Learning	0	12		7.2
Advantage	0	14		7.4
Insight	0	12		7.5
Impression	0	16		7.7
Knowledge	0	32		8.2
Experience	0	65		8.7

Figure 9. Comparison of subject collocation.

The third part of difference lies in the collocation with adverbs (see Figure 10). The adverb to describe *obtain* can be categorized into adverbs of manners (*dishonestly*, *illegally*, and *fraudulently*), adverbs of time (*subsequently*, *previously*), and adverbs of place (*elsewhere*). The common adverbs to describe *gain* are adverbs of degree (*steadily*, *substantially*, and *considerably*) and adverbs of frequency (*rapidly*, *gradually*). The findings correspond to the collocation difference in the part of noun collocation.

Collocation(adverb)	Frequency	Frequency	Collocation	Collocation
	of	of	Significance	Significance
	Obtain	Gain	of	of Gain (MI)
			Obtain(MI)	
Dishonestly	21	0	9.2	
Illegally	11	0	7.8	
free	9	0	7.8	
Fraudulently	6	0	7.5	
Readily	14	0	6.8	
Independently	6	0	6.6	
Subsequently	13	0	6.5	
elsewhere	5	0	5.8	
Previously	14	0	5.8	
Either	15	0	5.7	
Directly	11	0	5.5	
Successfully	0	8		5.8
Considerably	0	6		5.8
Rapidly	0	11		6.1
Gradually	0	9		6.1
Substantially	0	7		6.4
Steadily	0	8		6.7
Financially	0	7		7.3

Figure 10. Comparison of adverb collocation.

## The Difference in Semantic Prosody

Concordance lines in BNC of *Sketch Engine* will be used to analyze the semantic prosody of the two words. Again, the lemma of verb shall be chosen and KWIC shall be ticked. The words immediately after (to the right of) the selected words are in alphabetical order.

The Appendix 1 is "obtain" in 100 alphabetically sorted random concordances. It should be noted that as analyzed in the collocation part, the phrase "obtain (money/property) by pretences" as "v+prep+obj" pattern indicates negative connotation of criminal charge, though "obtain money/property" as "v+noun" pattern does not endow with the negative connotation. In this sense, the whole concordance will be analyzed in terms of the patterns with preposition.

In terms of pattern "v+prep+obj", the preposition by, which accounts for 17 concordances, shows the way to obtain, while the preposition from, which accounts for 11, explains the channel in which something been obtained. "Obtained by false pretences", "obtained by fraud", and "obtained a pecuniary advantage by deception" are all endowed with the negative connotation and appear four times in total. The other collocations, such as "authorization must be obtained from HMIP", "the oil obtained from the plant", and "material can be obtained by calling", do not have either positive or negative connotation.

Among the remaining 83 concordances, noun or noun phrase collocated with *obtain* will be the focus of analysis. *Abortion, problem*, and *low wage* as objects of verb refer apparently to bad things. *License, support, consent*, and *qualification* are endowed with positive denotation. The remaining 67 noun collocations, such as *result, figure*, and *treatment*, do not have either positive or negative denotation. Figure 11 demonstrates the types of semantic prosody, frequency, and corresponding percentage. To conclude, "*obtain*" has neutral or mixed semantic prosody.

Types of Semantic Prosody	Frequency	Percentage
Positive	24	24%
Neutral or ambiguous meaning	69	69%
Negative	7	7%

Figure 11. Semantic prosody of obtain.

The concordances of *gain* are also alphabetically sorted (see Appendix 2). It can be found that *gain* is dominantly collocated with the nouns which have positive denotation, such as *popularity*, *independence*, *confidence*, and *value*. Though "*impression*" and "*time*" are neutral noun, "correct impression" and "accurate time" as noun phrases have positive denotation. *Weight*, *height*, *information*, and *statistics* are neutral, but it should be stressed that *gain* means *increase* when collocate with *weight* and *height*. *Nothing* and *none* are ambiguous or with neutral denotation. The following diagram (see Figure 12) manifests the types of semantic prosody, frequency, and corresponding percentage. To conclude, *gain* has positive semantic prosody overall.

Types of Semantic Prosody	Frequency	Percentage
Positive	72	72%
Neutral or ambiguous meaning	28	28%
Negative	0	0%

Figure 12. Semantic prosody of gain.

## **Conclusion and Implication**

Based on the corpus analysis, the differences of two verb synonyms—obtain and gain lie in genre, colligation, collocation, and semantic prosody. Concrete noun is more collocated with obtain, which partly explains its more frequent usage in pure and practical science. The passive voice pattern with preposition is more widely used in obtain and the criminal charge like "obtain by deception" is often used in law and business English. Generally speaking, obtain has mixed or neutral semantic prosody. The verb form gain is more collocated with abstract noun and most of them endow with the positive semantic prosody. It is often used in the text types of commerce and economy, politics, and social science. These findings can have a profound impact in the fields of dictionary editing, English language teaching, and translation into other languages.

First, Oxford Dictionary failed to add the semantic prosody and provided the frequently used collocation such as "obtained by pretence", though it has made an attempt to introduce the collocation information into the definition. McGee (2012a) also claimed that the dictionaries, whether corpus-based or not, have not always noted important semantic prosody information related to words, especially difference in semantic prosody of synonyms.

Second, in language teaching, the traditional practice of explaining meanings to learners by offering synonyms should be used with caution since synonyms usually differ in their collocational behavior and semantic prosodies (XIAO & McEnery, 2006). In terms of data-driven learning, an inductive approach to language learning, learners are encouraged to become language researchers to explore corpus data and looking for answers to language questions, such as comparing synonyms (McGee, 2012b). But the learners' English level and purpose of English learning should be considered in organizing corresponding classroom activities. The study by Yeh, Liou, and LI (2010) in TsingHua University of Taiwan manifested that the possibility of using online materials and concordancing to increase EFL learner's awareness and application of synonymous adjectives. But students have to receive thorough training in induction skills before they study corpus data.

Third, this study also provides guidance for translating English into Chinese based on the language usage, particularly semantic prosody. Professor XIAO and McEnery (2006) undertook a cross-linguistic analysis of collocation and semantic prosody of near synonymy, drawing upon data from English and Chinese. Their contrastive analysis showed that semantic prosody and semantic preference are as observable in Chinese as they are in English. In this sense, to translate "obtain money by deception" into Chinese, it should be emphasized that *obtain* is not "得到" (*dedao*), but "骗取" (*pianqu*), which is endowed with the negative

semantic prosody. Similarly, gain in gain respective is translated as "赢得" (yingde), which has positive semantic prosody.

#### References

Biber, D., Conrad, S., & Reppen, R. (2006). *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.

Firth, J. (1957). Papers in linguistics. Oxford: Oxford University Press.

Francis, G., Hunston, S., & Manning, E. (1996). Collins cobuild grammar patterns 1: Verbs. London: Harper Collins.

Halliday, M. A. K., Teubert, W., & Cermakova, A. (2005). *Lexicology and corpus linguistics: An introduction*. London and New York: Continuum.

Hunston, S. (2002). Corpora in applied linguistics. Cambridge: Cambridge University Press.

McEnery, T., & Hardie, A. (2012). Corpus linguistics: Method, theory and practice. Cambridge: Cambridge University Press.

McEnery, T., XIAO, R., & Tono, Y. (2006). *Corpus-based language studies: An advanced resource book*. London and New York: Routledge.

McGee, I. (2012a). Collocation dictionaries as introductive learning resources in data-driven learning—An analysis and evaluation. *International Journal of Lexicography*, 25(3), 319-361.

McGee, I. (2012b). Should we teach semantic prosody awareness? RELC Journal, 43(2), 169-186.

Oxford Dictionary Online. (2015). Retrieved from http://www.oxforddictionaries.com

Sinclair, M, J. (1996). The search for units of meaning. Texus, 9(1), 75-106.

Sinclair, M, J. (2004). Trust the text: Language, corpus and discourse. London: Routledge.

Stubbs, M. (1996). Text and corpus linguistics. Oxford: Blackwell.

XIAO, R. Z. (2008). Theory-driven corpus research: Using corpora to inform aspect theory. In A. Ludeling and M. Kyto (Eds.), *Corpus linguistics: An international handbook* (pp. 987-1008). Birling: Mouton de Gruyter.

XIAO, R., & McEnery, T. (2006). Collocation, semantic prosody, and near synonymy: A cross-linguistic perspective. *Applied Linguistics*, 27(1), 103-129.

Yeh, Y., Liou, H., & LI, Y. (2007). Online synonym materials and concordancing for EFL college writing. *Computer Assisted Language Learning*, 20(2), 131-152

## Appendix 1

20% returns from assured tenancy schemes with non-recourse
4 lbs costing 60p and then said that 45 was halfway between
4 (3,225 by applying for a loan in he name of Garwand to
a favourable outcome , leave should be granted provided the
a licence from the Garwand to leave a licence from the Garwand to
a favourable outcome , leave should be granted provided the
a licence from the Garwand to leave the leav companies. 'Why should investors risk money when they could would cost: the pupil doubled 2 lbs coating 30b to 6,500. The jury found Yousefin ot guilty of attempting to 6,500. The jury found Yousefin ot guilty of attempting to 6,500. The jury found Yousefin the suity of attempting to 6,500. The jury found Yousefin the suity of attempting the suity of the obtain obtain obtained obtain obtain obtaining obtaining obtain obtain obtaining obtain obtained obtaining obtaining obtaining obtain obtain obtain obtain obtained obtained obtain obtained obtained obtaining obtained obtaining obtained obtain obtained obtained obtaining obtained obtain obtain obtained obtain obtained obtained obtain obtained obtained obtained obtain obtained obtained obtained obtained obtain obtaining obtaining obtain obtain obtained obtain obtaining obtain obtaining obtain obtaining obtained obtaining obtained throughout the ancient world — not only in the classical under high pressures . As far back as 1949 P. W. Bridgman up-to-date records of research in progress in the universities using isotopically substituted samples . As we have seen in using LISTREL or LIFESPAN Relational Database Interface (RDBI vitamin A from plant sources (such as carrots ) and must was an advantage . British flights to the Caribbean and South when their conditions hold , as those conditions are relaxed . with the necessary finance , this railway may well become the written confirmation from the interested party that it is in . Despite difficult market conditions we have been able to . This is not to set themselves up as nutritionists or as ; and there were fewer time constraints on its use. Although (deleted:formula) Comparison with (14) shows that we have (deleted:formula) . Measurements of the Doppler shift at obtained obtained obtaining obtained obtained obtain obtained obtaining obtained obtain obtained obtained obtained obtained obtain obtain

## Appendix 2

Sp to 4 , after an interim rise from 15.4m to 26.1m a clearer understanding of the plan and construction (190 ) .

a confidence-boosting victory in the Marsh Green Haiden By a deserved reputation for looking after its public interests .

a qualification which recognises competence in the worplace.

a raptuation for not being over-reliable, perhaps he suffered a ridge . Make you way to the northwest end of Lochan a SOJCERT Part 1 , it was suggested that students would normally access to any part of the Town Hall. This is proving access to any part of the Town Hall. This is proving access to it increasingly became seen as problematic. The them a society of the company of the a S% midway rise to Pounds 141m pre-tax . Eastern Electricity the brilliance it once had . From the gallery , one he has his easiest task to date at Lingfield and can to compel its own LEA to provide sufficient accommodation will challenging new route whereby those working in the industry can more in colour, where, for example , many actors did not it meant that, for a while , he around two miles on the loch shore then turn right to and III should be amalgamated with SCOTCERT . In order to were replaced by ramps so that now wheelchair users can of the period . What Renaissance culture was and how we have phoned; some may have 'visited' but not limits for women , but in those areas where women have or professional placement in their third year , thereby forming outside the test centres in Beijing just in order to . The adventurers must each make a successful Cl test or insight into the qualities of the Namemanagement and for them to to make to stay in business — and to grow . six shirt by the end of that 1977-78 season . After Certificate Programme , including the Case Study and who have seemed speaking when you felt like it was a way to of full-term babies found that more than three-quarters sport than the others, cricket has played a part in pardons dissipated the public support which the government had company itself . Surveys have shown that the added value depth of learning and the sense of satisfaction and achievement so. Kurt Walcheim elected UN General Secretary . Idi Amin cent of the country's mineral exports , is trying to the Tories' 57, believes it needs another 25 to the silly walks , deceased parrot and other images that have system built for the 1950s . Another technological phrase that flurst , near Richmond , and Kate , from Huworth , him , punishing him , because after all , far from take place , for Cnut was shrewed enough to appreciate that already full to overflowing with the crowds who had failed to pay out the money if Addamax loss — it will not 'Well , now you have . What have you . Of co gain gain gain gain gained gain gain gain gaining gain gain Gaining gaining gained gain gained gaining gained gained gained gains gain gain gained gained gained gaining gaining already full to overflowing with the crowds who had failed to pay out the money if Addamax loses — it will not "Well, now you have . What have you . Of course, the moral or symbolic stature that he and advertising are a continuous process, and knowledge . It examines whether the shareholders of the acquiring firms members and wider public, in their tens of millions, — "Why am I doing this?" The experiences visited two schools with project experience in order to the good business they did. Dealers at Harvard Securities who on his engines, which demonstrated their high efficiency and Gowing and Victor Pasmore at Newcastle University, where he were anointed by their respective bishops. Thus when Charles the dodur becomes very bad indeed . Certainly nothing would be simple mathematical correlation between the villus height and an attempt to snatch him back. But he is the lion's share of growth. However, mail-shots are of a rooted sorrow. Unlike her husband who has happened to market share? Well we 've er we 've years in Northern Rhodesia, but some years after the country and given an opinion as to its value in order to ombudsman and the Faculty of Advocates, he has failed to returning to the crouched position. It is essential that you may be a ruse by which low-ranking males are able to English-language instruction. The Buddhist revival to include me? The me of 1989 exists and can. Against all predictions, the Nationalists lost two seats and control, nor held responsible for unfair gains when they have apparently happily married man with no enemies. He stood to It is in Raverna that the most accurate impression can be run for it. As he slithers across the mul I slage trailer. Wherever it's working, the JFT soon only for the one-day internationale. Ruscell supporters occurses are cut out in joint schemes, and the students Hurn, chairman and chief executive, said the success at and thereby derive contextual effects which would not have been of the welfare state that everybody else in this room grace and the support o gained gained gained gain gain gain gained gain gained gained gained gained gained gain gaining gaining gained gained gained gain gain gain gain gained gain gained gained gain gained gain gains gain gain gaining gained gained gained gain gains gained gain gain gaining gained gain gaining gain gained gain gain gaining gain gain gaining gain gain gain gained gain gained gained gain gained