

# A Corpus-Based Comparative Study of English Synonyms: Taking *Improve, Promote, and Strengthen* as an Example

Yaru Liu<sup>1\*</sup><sup>1</sup>Southwestern University of Finance and Economics, ChinaDOI: [10.36348/sijll.2023.v06i09.005](https://doi.org/10.36348/sijll.2023.v06i09.005)

| Received: 18.08.2023 | Accepted: 25.09.2023 | Published: 29.09.2023

\*Corresponding author: Yaru Liu

Southwestern University of Finance and Economics, China

## Abstract

As synonyms account for more than 60% of English vocabulary, improving the ability to distinguish synonyms will help speakers produce language in a more precise and efficient manner. This paper adopts a corpus-based approach to explore the usage of synonymous words in terms of language register, collocates and semantic prosody, exploring how these synonyms are used by native speakers and hoping to offer some guidance to English learners and translators.

**Keywords:** Corpus; synonym; collocates; register.

**Copyright © 2023 The Author(s):** This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY-NC 4.0) which permits unrestricted use, distribution, and reproduction in any medium for non-commercial use provided the original author and source are credited.

## 1. INTRODUCTION

Synonyms, which account for more than 60% of English vocabulary (He, 2003), have always been the focus and a challenge in second language learning and usage. Being highly similar in form, meaning and usage, they are prone to be misused by non-native speaker, consequently causing lots of communication problems. Improving the ability to distinguish synonyms will help speakers produce language in a more precise and efficient manner.

By taking the BNC corpus as the research tool, this paper discusses the slight but significant differences of “improve”, “promote” and “strengthen” in practical usage. With the similar meaning of increase and enhancing, they appear indistinguishable to most nonnative English learners. And according to the research conducted by Yu and Lu (2022), these three words are among the most frequent verbs of the English version of the government work report, indicating that the discrimination of these three words is of great significance to both language learning and the translation of national publicity materials.

## 2. LITERATURE REVIEW

Synonyms can be distinguished in many ways, and the traditional one depends on intuition, empirically analyzing the words’ meaning and matching them to different Chinese equivalents. Nevertheless, such method has neglected the real usage of words, leading to the unavoidable misuse of synonyms in specific

contexts. Gradually scholars have adopted corpus method in figuring out the differences between words with like meaning.

Corpus refers to a large collection of well-sampled and processed electronic texts. With the aid of computer tools, scholars are able to differentiate words by exploring their natural occurrences. Yang and Liu have put forward some corpus-based approaches to the differentiation of English synonyms. Their ideas, including examining the distribution difference in different language registers, retrieving the significant collocations, and exploring the differences in terms of colligation and semantic prosody by observing the synonymous words’ collocation features (2005), have inspired the latter writers (Weng, 2020; Wen, He & Zhang, 2007; Ren, 2008) in their studies of synonyms.

Firstly, at the macro level, linguists have paid great attention on language register when exploring the differences of words. Halliday defines register as “a variety of language, corresponding to a variety of situation” and “clustering of semantic features according to situation type” (Halliday, 1985). According to *Longman Dictionary of Language Teaching and Applied Linguistics*, “a particular register often distinguishes itself from other registers by having a number of distinctive words, by using words or phrases in a particular way, and sometimes by special grammatical constructions”(2000). So, the comparison of the register distribution of synonyms is instructive in deciding which

word should be selected in various spoken or written contexts.

Secondly, the micro perspective of investigating synonyms is the examination of their collocates, colligation and semantic prosody. And the study of the collocation of the words constitutes the basis of this level.

Some earlier scholars have taken collocation as being partially or totally arbitrary (Lewis, 1993, 1997, 2000; Nation, 2001; Woolard, 2000; McCarthy, 1990). Lewis, one of the most influential Americans applied linguists in the field of English language teaching, advocates that collocation is not determined by logic or frequency, but is arbitrary (Lewis, 1997). And to some extent it cannot be predicted or reasonably explained (Mackin 1978). Consequently, language learners are forced to learn and memorize collocations as a whole unit without any explanation or any other help to understand them.

More recent studies have lodged an objection to the opinion of the arbitrariness of collocations, saying that “collocation is not simply an arbitrary phenomenon but is a process which can be partially explained” (Walker, 2011), and students can begin to perceive why one word frequently occurs in the company of another (Crowther, Dignen, & Lea, 2002; McCarthy & O’Dell, 2005). Gradually techniques drawn from corpus linguistics and cognitive linguistics have been combined to study factors that influence the way in which collocations are formed. Corpus linguistics can provide a more objective perspective and the possibility to analyzing a massive text for cognitive study, while cognitive linguistics contributes to collocation research by offering theoretical guidance. Liu (2020) is one of the studies which critically examined the accepted definition of collocation in this way.

With an increasing number of linguists rejecting the arbitrariness of collocations, studies on colligation and semantic prosody of words by investigating their collocates have emerged. Colligation can be seen as the grammatical collocation which suggests the inter-relation of grammatical categories in syntactic structure (Firth, 1957). Lewis differentiates colligation from collocation in a simple way, saying that collocation refers to the way in which a word cooccurs with another word, while colligation is the way that a word cooccurs with a particular grammatical pattern (2000).

And semantic prosody was coined by Louw in an article published in 1993. He defined it as “a

consistent aura of meaning with which a form is imbued by its collocates” (1993, p. 157). Generally semantic prosody is classified into three kinds: positive, neutral and negative. As Louw put it, semantic prosody is the result of a diachronic process in which meaning has been transferred from one word or words to another, and the identification of a positive or negative prosody of a lexical item relies on whether it frequently occurs with collocates which refer to desirable or undesirable items or events.

Based on previous studies, Walker concludes that “the process of collocation is influenced by, for example, the precise meaning or meanings of a particular lexical item, the use of metaphor, and any phraseological behaviour or semantic prosody associated with the item”.

### 3. METHODOLOGY

The British National Corpus (BNC) is a 100-million-word collection of electronic texts from a wide range of genres, including spoken, fiction, academic, magazines and newspapers. It was designed to represent the British English usage in the later part of the 20th century. The corpus analysis tool used here is a service provided by Lancaster University, which can be freely accessed at <http://bncweb.lancs.ac.uk/>.

The detailed research steps are as follow:

1. Carry out a lemma search for the verbs “improve”, “promote” and “strengthen”, then use the Distribution function to get their frequency in different language registers.
2. Conduct the lemma search again, use the Collocation function, and change the Collocation window span: to ‘1 Right - 1 Right’, and under Filter collocates by tag specify ‘any noun’. The list of their most frequent noun collocates ranked by log-likelihood will be generated.

### 4. RESULTS AND DISCUSSION

#### 4.1 Register Analysis

Language register is defined as the linguistic variation which describes the complicated ways speakers modify their language use on the basis of social cues, communication context, and personal expression. It indicates speaker-audience relationship and linguistic style such as level of formality. The comparison of the register distribution of synonyms is instructive in deciding which word should be selected in various spoken or written contexts. Table 1 Shows the distribution of these three verbs in registers of different levels of formality.

**Table 1: The distribution of "improve", "promote" and "strengthen"**

Text types		Verbs		improve		promote		strengthen	
		Freq.		Freq.	%	Freq.	%	Freq.	%
Spoken		647		274	5.9	4.2		67	2.3
Written	Academic prose	1878		1445	17.2	22.1		624	21.4
	Fiction and verse	342		116	3.1	1.8		122	4.2
	Newspapers	1385		655	12.7	10.0		337	11.5
	Other written material	6680		4045	61.1	61.9		1771	60.6
	<b>total</b>	<b>10285</b>		<b>6261</b>	<b>94.1</b>	<b>95.8</b>		<b>2854</b>	<b>97.7</b>

As the data suggests, "improve" is more extensively used by native speakers as it occurs far more frequently than the other two in all text types and especially in spoken and newspaper domain. All the three verbs have more occurrences written text than in the spoken language, but relatively "strengthen" occurs less frequently compared with "improve" and "promote" in the spoken context because its frequency in the spoken is less than one fifth of that in the written text, indicating the casual usage of "strengthen" in spoken materials is less acceptable. And by comparing the ratio of frequency in academic prose to that in whole written material, it can

be concluded that "promote" is prone to be used in academic domain.

**4.2 COLLOCATION ANALYSIS**

**4.2.1 Overview of the Noun Collocates**

Though the three selected verbs have similar meaning, in real practice there are still slight but significant differences that can be extracted by analyzing their collocations.

Table 2 was generated through the collocation tool with changing the Collocation window span to 1 Right-1 Right, and the result are filtered by "any noun" and ranked by Log-likelihood.

**Table 2: Noun Collocates of "improve", "promote" and "strengthen"**

No.	improve	Freq.	LL	promote	Freq.	LL	strengthen	Freq.	LL
1	efficiency	56	456.39	homosexuality	17	176.48	links	22	197.77
2	performance	71	434.56	competition	29	173.76	Gloucester	6	38.85
3	matters	43	252.45	equality	17	145.22	consumer	6	35.89
4	productivity	31	255.58	growth	23	113.43	France	7	30.01
5	standards	42	238.33	co-operation	16	106.95	relationships	5	25.14
6	safety	39	224.49	awareness	14	90.70	democracy	4	21.11
7	relations	41	218.46	health	23	84.99	cooperation	3	21.05
8	communication	26	174.91	peace	16	79.76	exercises	3	19.72
9	profitability	18	158.01	efficiency	10	57.45	relations	5	19.07
10	access	31	155.58	research	18	57.12	bonds	3	18.64

Table 2 reflects that "improve" is always followed by nouns which can be measured by numerical value, such as efficiency, performance, productivity and profitability. In its meaning there is an emphasis on performing better in some aspects or making things like performance, safety, relation more satisfactory. And "promote" is likely to occur in company with words with

more macro meaning, such as equality, health and peace. Different with improve, its significant noun collocates usually cannot be measured by value, and its prominent meanings concluded from the concordances are endorsing something and advancing or furthering the process of something, as can be seen in the examples below:

prohibits authorities from intentionally **promoting** **homosexuality** or publishing material  
*to endorse*

It is also appointing a coordinator to **promote** **research** on electro-active polymers  
*to further the process*

Its aim was to **promote** **peace** and international understanding  
*to further the process*

The collocates of "strengthen" indicate its subtle implication, that is, to make something stronger or more durable. And the verb's objects are usually the

links between things such as relationship, cooperation, relation and bond. The following concordance lines serve to illustrate this point.

This may have **strengthened Gloucester's hand** in the region  
*to make something stronger*  
 consensus on the need to **strengthen consumer representation**  
*to make something stronger*  
 his government's commitment to **strengthening relations** between the United Kingdom  
*to make something stronger/more durable*

Nevertheless, in table 2 there are some shared noun collocates such as “efficiency” (improve efficiency, promote efficiency), “relations” (improve relations, strengthen relations) and “cooperation” (promote cooperation, strengthen cooperation). It must be noticed that no absolute bound exists between these three synonyms as collocations are viewed more in terms of probability. According to Halliday, “the native speaker’s knowledge of his language will not take the form of his accepting or rejecting a given collocation: he will react to something as more acceptable or less acceptable on a scale of acceptability” (1966, p.159). This means that there are virtually “no impossible collocations, but that some collocations are much more likely to occur than others.” (Walker 2011, p.293). From

this perspective, it can be said that “improve efficiency” is more acceptable than “promote efficiency”, and it is more native to say “strengthen links” rather than “improve links”. Grabbing the slight but significant differences between synonyms therefore is helpful for language learners and translators.

**4.2.2 Evaluative Tendencies of the Collocates**

In this part, I retrieve 50 most significant noun collocates of “improve”, “promote” and “strengthen” to examining the collocational behavior of these three verbs. Some of the collocates have a positive atmosphere while others are neutral or negative, as can be seen from the following typical examples.

programme of the same name will **promote growth** in mutual understanding  
*Positive (desirable and pleasant)*  
 it contains two clues for ways to **improve things.**  
*Neutral*  
 modernise its nuclear capability to **strengthening deterrence.**  
*Negative (undesirable)*

The classification of the collocates are showed in the following tables.

**Table 3: The classification of the collocates of "improve"**

<b>Collocates with the positive atmosphere: 18 (36%)</b>
efficiency, productivity, safety, profitability, access, quality, health, survival, circulation, morale, understanding, education, stability, fitness, comfort, recognition, security, energy
<b>Collocates with the negative atmosphere: 1 (2%)</b>
patient
<b>Collocates with the neutral atmosphere: 31 (62%)</b>
relations, communication, fuel, matters, standards, communications, services, conditions, working, facilities, things, links, relationships, transport, living, customer, housing, staff, service, product, road, prison, training, pupils, Britain, women, treatment, management, data, labour, performance

**Table 4: The classification of the collocates of "promote"**

<b>Collocates with the positive atmosphere: 27 (54%)</b>
competition, equality, growth, awareness, health, peace, efficiency, democracy, healing, harmony, independence, cooperation, innovation, stability, opportunities, understanding, reform, learning, freedom, education, policies, co-operation, energy, improvements, symmetry, fertility, adjustment
<b>Collocates with the negative atmosphere: 1 (2%)</b>
homosexuality
<b>Collocates with the neutral atmosphere: 22 (44%)</b>
tourism, trade, sales, exports, membership, employment, investment, research, enterprise, medau, discussion,staff, community, captain, contact, training, legislation, consumer, change, measures, debate, women

**Table 5: The classification of the collocates of "strengthen"**

<b>Collocates with the positive atmosphere: 15 (30%)</b>
democracy, cooperation, confidence, morale, accountability, motivation, competition, co-operation, unity, recovery, support, security, family, education, perestroika
<b>Collocates with the negative atmosphere: 1 (2%)</b>
deterrence
<b>Collocates with the neutral atmosphere: 34 (68%)</b>
relations, brand, communication, student, Gloucester, France, consumer, America, Germany, Britain, community, countryside, teeth, leg, branch, management, parliament, measures, channels, capitalism, Canada, claims, relationships, bonds, devices, executive, links, csce, gen, copyright, exercises, defences, protection, arguments

The data shows that none of these three verbs tend to occur with negative collocates. As “improve” and “strengthen” have respectively 62% and 68% of collocates being neutral, it is reasonable to say that they are in strong association with words which have a neutral atmosphere. Compared with “improve” and “strengthen”, “promote” has a greater proportion of positive collocates, so it is more likely to occur in accompany with nouns which refer to desirable events or items.

## 5. CONCLUSIONS

In conclusion, “improve”, “promote” and “strengthen” have similar meaning but are slightly different in usage. “Improve” is more frequently and extensively used by native speakers than the other two. “Promote” is tend to occur in academic text while “strengthen” rarely occurs in spoken language. In the respect of collocation, “improve” always precedes nouns which can be measured by numerical value and “strengthen” often occurs in accompany with nouns which has a meaning of “link” between objects. And the distinct meaning of “promote” is endorsing something or advancing or furthering the process of something. Additionally, “promote” is more likely to occur in accompany with nouns which refer to desirable events or items.

## REFERENCES

- Crowther, J., Dignen, S., & Lea, D. (Eds.). (2002). Oxford collocations dictionary for students of English. Oxford, England: *Oxford University Press*.
- Firth, J. R. (1957). Modes of meaning. In F. Palmer (Ed.), *Papers in linguistics, 1934-1951* (pp. 190-215). Oxford, UK: *Oxford University Press*.
- Halliday, M. A. K. (1966). Lexis as a linguistic level. In C. Bazell, J. Catford, M. A. K. Halliday, & R. Robins (Eds.), *In memory of J. R. Firth* (pp. 148-162). London, England: *Longman*.
- Halliday, M. A. K. (1985). *An Introduction to Functional Grammar*. London: *Edward Arnold*.
- Lewis, M. (1993). *The lexical approach* (Vol. 1, p. 993). Hove: Language teaching publications.
- Lewis, M. (1997). *Implementing the lexical approach*. Hove, England: Language Teaching Publications.
- Lewis, M. (2000). *Teaching collocations*. Hove, England: Language Teaching Publications.
- Liu, D. (2010). Going beyond patterns: Involving cognitive analysis in the learning of collocations. *TESOL quarterly*, 44(1), 4-30.
- McCarthy, M. (1990). *Vocabulary*. Oxford University Press.
- McCarthy, M., & O'Dell, F. (2005). *English Collocations in Use. Intermediate*. Cambridge: *Cambridge University Press*.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge, England: Cambridge University Press.
- Siepmann, D. (2005). Collocation, colligation and encoding dictionaries. Part I: Lexicological aspects. *International Journal of Lexicography*, 18(4), 409-443.
- Walker, C. P. (2011). A corpus-based study of the linguistic features and processes which influence the way collocations are formed: Some implications for the learning of collocations. *Tesol Quarterly*, 45(2), 291-312.
- Woolard, G. (2000). Collocation-encouraging learner independence. *Teaching collocation: Further developments in the lexical approach*, 28-46.
- 任培红.(2008).近义词的教与学:以 Common 和 Ordinary 为例. *解放军外国语学院学报*(04),57-60.
- 喻岩 & 鲁修红.(2022).基于《政府工作报告》英译本的中国形象研究. *湖北工业大学学报*(03),74-78.
- 张继东,刘萍.(2005).基于语料库同义词辨析的一般方法. *解放军外国语学院学报*(06).
- 温玲霞,何明刚 & 张吉吉.(2007).基于语料库数据的近义词语义韵调查——以 rather,fairly 为例. *沈阳大学学报*(05),56-60.
- 翁菁岭.(2020).基于 COCA 语料库的近义副词差异性研究——以 however、nevertheless 和 nonetheless 为例. *才智*(28),37-39.