

投稿類別：英文寫作

篇名：

Lexical Use Differences Between British English and American English: A Corpus-Based Study

作者：

印心瑜。台北市立景美女中。高三善班

指導老師：

呂翠屏老師、王伯雅老師

I. Introduction

As a long-standing joke goes, the United Kingdom and the United States are “two countries divided by a common language.” English is the world’s most commonly used language in the 21st century, but there are a variety of dialects. Among them, American English and British English are two major dialects of English that can be really confusing, especially when it comes to lexical uses. They seem similar but still differ in a considerable number of ways.

British English is the English language spoken and written in Britain or, more broadly, throughout the British Isles. According to Tom McArthur in the *Oxford Guide to World English*, British English shares “all the ambiguities and tensions in the word *British* and as a result can be used and interpreted in two ways, more broadly or more narrowly, within a range of blurring and ambiguity.” On the other hand, the use of English in the United States is a result of British colonization. The first wave of English-speaking settlers arrived in North America during the 17th century, followed by further migrations in the 18th and 19th centuries. Since then, American English has been influenced by the languages of German, Dutch, Irish, Spanish, and other languages of immigrants coming to the United States in successive waves.

Over the past 400 years, the form of the language used in the United States and that used in the United Kingdom have diverged in a few minor ways. Differences between the two include pronunciation, grammar, vocabulary (lexis), spelling, punctuation, idioms, and formatting of dates and numbers. The differences in the grammar structure, no matter written or spoken, tend to be much fewer than those of other aspects of the language in terms of mutual intelligibility. By contrast, there are a small number of words that have completely different meanings in the two dialects. Some of them are even not used in the other dialect.

A well-known example is the lexical difference of *football* in the United Kingdom versus *soccer* in the United States, when it comes to the ball-kicking sports. Confusingly, *football* in the US is also a very common sport where teams throw a *football* to each other trying to score a touchdown, *while football* in the UK is, literally, closer to our intuition, since the British *football* refers to the ball game in which the players can only kick the ball with feet.

This research paper aims at analyzing the differences of lexical use between American English and British English, and at discovering the possible cause of the differences through corpus research.

II. Thesis

A. Study 1: Word Pair Categorization

1. Method:

In this paper, we first selected 150 pairs of lexicons that share the same meaning in American English and in British. These vocabulary pairs were drawn from *The Secret Life of Words: How English Became English*, *A History Of The English Language*, and the Internet. Then, the words pairs were assorted into different categories based on the aspects the words refer to.

2. Result and Discussion:

The categories listed include Food, Clothes, Building, etc. It is obvious that these differences are mostly associated with daily life encounters. For example, when it comes to clothes, there are many differences found between American English and British English. The Americans say *hat*, *bathrobe*, *sweater*, *diaper*, and *sneakers*, while the British use, respectively, *bonnet*, *dressng gown*, *jumper*, *nappy*, and *trainers*.

Some extreme, even crazily confusable, pairs exist. For instance, a British *vest* is an American *undershirt*, and in the meanwhile, an American *vest* is a British *waistcoat*. Thus, going shopping and travelling around can be really confusing in the two countries if the visitor is not familiar with or not aware of such lexical differences. For example, the British go to the *chemist's* while the Americans go to a *drugstore* or *pharmacy* to buy medicine or body-care products. The British would get a *booking* at a *pub*, while the Americans would have a *reservation* at a *bar*.

The lexical difference may result from the lost of a particular word in one of the dialect. For example, some older English words like *fortnight*, meaning “two weeks”, have fallen out of use in the USA, but are still in common use in Britain.

Though American English seems to be a new branch of dialect than British English, an American lexicon may be the one that originates from the older form of English. For instance, the American word *fall* predates the British *autumn*. The British *autumn* became the favored British English word after the American civil war, when French was highly fashionable among British speakers. The USA kept the older word *fall*. Most people would say that *autumn* sounds more formal than *fall*. The same reason is true with pairs such as American *cotton candy* versus British *candy floss*, and American *faucet* versus British *tap*.

Sometimes, the difference of lexical use occurs because a dialect found a need to use one more distinctive term to describe the subtle sub-sense while the other dialect did not distinguish such a subtle distinction. Take the word pair *biscuit* versus *cookie* for example. According to the *Oxford English Dictionary (OED)*, the word *biscuit* comes originally from the Latin *biscotum (panem)*, which means bread ‘twice baked.’ This would explain the hard, crunchy quality of a British *biscuit*. The word *cookie* opens up a whole other can of worms. In the UK, a cookie is a soft, squishy, moist biscuit (for lack of a better word). British *cookies* tend to be bigger and more substantial than British *biscuits*. The word *cookie* comes from the Dutch *koekje*, meaning ‘little cake,’ and could have been popularized in the US due to early Dutch colonization. In the US, a *cookie* covers both what the British would call a *biscuit* and a *cookie*.

B. Study 2: Word Frequency Analysis of the Word Pairs

1. Method:

The lexical pairs were further investigated by looking up their frequencies in the Corpus of Contemporary American English (COCA) and The British National Corpus (BNC). These two corpora respectively represent the use of American English and that of British English. Thus, the use frequency of a word in the COCA signifies how often the word is used in American English, while the use frequency of the same word in the BNC stands for how frequently the word is used in British English.

The Corpus of Contemporary American English (COCA), created by Mark Davies and Brigham Young University, is the largest corpus of American English currently available. It

appeared on the scene in 2008 and it includes around 450 million word. It is freely searchable and is the only publicly available corpus of American English to contain a wide array of texts from a number of genres. (Corpus Linguistics and the Description of English, pp. 170)

The British National Corpus (BNC) is the result of a major commercial and academic project, which started in 1991 and was run by the Oxford University Press, the Longman Chambers, the British Library, the Oxford University and the Lancaster University. It is a 100-million-word text corpus with samples of written and spoken British English from a wide range of sources. The corpus covers British English of the late 20th century from a wide variety of genres with the intention that it be a representative sample of spoken and written British English of that time.

Since the two corpora has a different size in words, it is not workable to directly compare the word pairs by the frequency directly derived from them. Thus, to compare their frequency with the same unit, the original frequency was converted into “per million frequency,” meaning the occurrence of a word out of every million words in the corpus. (Corpus Linguistics and the Description of English, pp.170) Sketch Engine, containing 400 ready-to-use corpora in more than 85 languages, was used to search for words in the BNC since it provides the per million frequency outcomes of the BNC. On the other hand, the scale of COCA is transformed into per million with the help of the software Excel.

2. Results and Discussion:

Table 1 below shows typical examples. 77 out of the 150 pairs investigated turned out to be this type. The first column is British English lexicon while the second column is American English lexicon.

Table 1. Typical Pairs (Unit: per million)

BE	AE	BE in BNC		AE in BNC	BE in COCA		AE in COCA
pavement	sidewalk	15.33	>	0.89	7.66	<	14.50
rubbish bin	trashcan	0.49	>	0.08	0.02	<	0.24

Take *pavement* versus *sidewalk* for example. “BE in BNC” refers to the use frequency of the BE

lexicon *pavement* in the BNC, the British English corpus, while “AE in BNC” stands for the use frequency of the AE lexicon *sidewalk* in BNC. “BE in COCA” is the use frequency of the BE word *pavement* in the COCA, the American English corpus, as “AE in COCA” means the use frequency of the AE word *sidewalk* in the COCA.

They serve as typical examples because the British English words reasonably show a tendency of higher frequencies than the American English words in the BNC, the British English corpus. On the other hand, the American English words used by Americans tend to have higher frequencies in COCA, the American English corpus, than their British counterparts in the COCA. As shown in Table 1, *pavement* is more frequent than *sidewalk* in the BNC. In the COCA, it is the other way around.

However, around half of the pairs (73 pairs) do not belong to the typical type (Type 1). As shown in Table 2, there are 3 other minor types. One example pair was provided for each type. For Type 2 the use frequencies of the word pairs in one corpus shows a reverse result, contradictory to the typical tendency, while the use frequency trend in the other corpus is typical. For example, *braces* is a British word but it is used widely in the US. Type 3 is a curious type. The use frequencies of a word pair is relatively the same in one corpus while those in the other corpus display a typical pattern. The phenomenon about the use of *cookie* versus *biscuit* discussed in Study 1 may be the strong support for this frequency result.

For Type 4, the use frequencies of the word pairs show opposite patterns to the typical type in both corpora, which is really rare. . *Tailback* is listed in the British dictionary, while *traffic jam* is in an American word. However, the frequency of *tailback* is high in COCA instead of NBC, and *traffic jam* has a higher frequency in NBC instead of COCA.

Table 2. Minor Types (Unit: per million)

	BE	AE	BE in BNC		AE in BNC	BE in COCA		AE in COCA
Type 2	braces	suspender	1.64	>	0.25	2.88	>	0.06
Type 3	biscuit	cookie	0.91	<	1.29	1.94	<	10.33
Type 4	tailback	Traffic jam	0.36	<	1.39	1.47	>	0.83

III. Conclusion

A reason why American English and British English can be so easily confused is because, when the people from Europe immigrated to America, they didn't want to speak like their colonial mother country or write like them. The people wanted to be themselves. They hoped that every physical object can be different.

There are times that their frequencies make them even more confusing. The reason may be that the Americans and the British have become closer along with the invention of advanced transportation and the Internet. Americans and British have more chances to communicate and to come into connection. Globalization is enhancing the fusion of the two dialects. Language changes through people's use. Further investigation can be conducted to discover interesting facts behind the use frequency of lexical differences among dialects.

IV. Reference

A. Books

1. Hitchings, Henry. (2012). *The Secret Life of Words: How English Became English*. (林俊宏, trans.) Common Master Press. (Original work published 2009).
2. Baugh, Albert. (2002). *A History of the English Language*. Routledge.
3. Hans Lindquist. (2009). *Corpus Linguistics and the Description of English*. Edinburgh University Press

B. Websites

1. Oxford Dictionary. <http://www.oxforddictionaries.com/words/british-and-american-terms>
2. Mandy Barrow. (2013). British Life and Culture. <http://resources.woodlands-junior.kent.sch.uk/customs/questions/americanbritish.html>
3. M. Boyanova. <http://www.studyenglishtoday.net/british-american-english.html>.
4. Cic Quarry. <http://www.vivquarry.com/wkshts/usavuk.html>.

C. Other Thesis

1. Guiseppian Scotto Di Carlo. (2013). Lexical Differences between American and British English:

A Survey Study.

2. Debjani Sanyal. (2013) A comparative Analysis of Lexical Variation in American and British English with Special Reference to A Few Selected words.