Semiconductor Optoelectronics

A COMPARATIVE ANALYSIS OF TOLERATE AND PUT UP WITH IN THE COCA

Namkil Kang
Far East University, South Korea

Abstract: The ultimate goal of this paper is to investigate the similarity between *tolerate* and *put up with* in the Corpus of Contemporary American English (COCA). For this goal, we compared *tolerate* and *put up with* by using three tests. A major point to note is two types (*tolerate* and *put up with*) reveal the same characteristic in two genres, whereas they do not in six genres. More specifically, *tolerate* is 25% the same as *put up with* in the genre analysis. A further point to note is that in the TV/movie genre, the Euclidean distance between *tolerate* and *put up with* is 11.16, which is the highest among eight genres. It amounts to saying that two types (*tolerate* and *put up with*) have the lowest similarity in the TV/movie genre. It is worthwhile noting, on the other hand, that the type *tolerate* is the nearest to the type *put up with* in the web genre. To be more specific, in the web genre, the Euclidean distance between *tolerate* and *put up with* is 0.37, which is lowest. From this it seems evident that *tolerate* and *put up with* have the highest similarity in the web genre. This paper shows that *tolerate people* and *put up people* have the highest frequency (32 tokens vs. 23 tokens). This paper further shows that the nouns *people*, *things*, *abuse*, and *harassment* are the collocations of both *tolerate* and *put up with*. It can thus be concluded that *tolerate* and *put up with* exhibit a very low similarity in the collocation analysis (the top 25).
Keywords: COCA, tolerate, put up with, token, genre, Euclidean distance, NetMiner,

1. Introduction

The main purpose of this paper is to compare *tolerate* and *put up with* in the Corpus of Contemporary American English (COCA). We aim to observe how similar the former and the latter can be in the COCA. We compare two types (*tolerate* and *put up with*) by doing three tests in the COCA. First, we try to compare two types in terms of a ranking analysis in the eight genres of the COCA. The so-called eight genres consist of the COCA and each genre shows the frequency of *tolerate* and *put up with*. In each genre, we can see which type is the preferable one for Americans. Second, we compare two types in terms of the so-called Euclidean distance in each genre. We calculate the Euclidean distance between two types in each genre and compare them. By doing this, we can see how similar two types can be in each genre. More than anything else, we can measure the nearest distance and the longest distance between two types. Third, we aim at comparing two types by observing their collocations. We can see how close they can be in the COCA. However, our list is limited to the top 25. Additionally, we Attempt to capture the similarity between two types in terms of the software package NetMiner. The NetMiner provides the network of two types and it captures the similarity between them.

2. The COCA

2.1. Eight genres

In what follows, we aim at comparing *tolerate* and *put up with* in the COCA. Table 1 shows the use of *tolerate* and that of *put up with* in the COCA:

Table 1 Use of tolerate and put up with

GENRE	All	BLOG	WEB	TV/M	SPOK	FIC	MAG	NEWS	ACAD
Tolerate	8,113	1,313	1,146	712	946	870	1,249	816	1,061
Put up with	4,596	776	666	916	570	787	398	364	119

It is important to note that the overall frequency of *tolerate* is 8,113 tokens, whereas that of *put up with* is 4,596 tokens. This in turn indicates that the type *tolerate* (8,113 tokens) is preferred over the type *put up with* (4,596 tokens) by Americans. From this it seems clear that Americans prefer using *tolerate* (8,113 tokens) to using *put up with* (4,596 tokens).

It is worthwhile noting that *tolerate* ranks first (1,313 tokens) in the blog genre, whereas *put up with* ranks first (916 tokens) in the TV/movie genre. As expected, the types *tolerate* and *put up with* show no similarity in rank-one. Note that in the blog genre, the number of *tolerate* is much higher (1,313 tokens) than that of *put up with* (776 tokens). This implies that *tolerate* is favored over *put up with* in the blog genre. Put differently, American bloggers prefer to use *tolerate* (1,313 tokens) rather than use *put up with* (776 tokens). It should be noted, on the other hand, that in the TV/movie genre, the use of *put up with* is by far higher (916 tokens) than that of *tolerate* (712 tokens). Clearly, American celebs prefer using *put up with* (916 tokens) to using *tolerate* (712 tokens).

It is worth mentioning that *tolerate* ranks second (1,249 tokens) in the magazine genre, whereas *put up with* ranks second (787 tokens) in the fiction genre. Again, the type *tolerate* does not pattern with the type *put up with* (in rank-two), hence no similarity. In the magazine genre, the use of *tolerate* is far higher (1,249 tokens) than that of *put up with* (398 tokens), which in turn indicates that American journalists prefer to use *tolerate* (1,249 tokens) rather than use *put up with* (398 tokens). It must be noted, on the other hand, that in the fiction genre, the number of the type *tolerate* is slightly higher (870 tokens) than that of the type *put up with* (787 tokens). It can thus be inferred that American writers are fond of using *tolerate* and *put up with* equally.

It is worth observing that *tolerate* ranks third (1,146 tokens) in the web genre, whereas *put up with* ranks third (776 tokens) in the blog genre. Again, *put up with* does not pattern with *tolerate* in rank-three, thus showing no similarity. It should be pointed out that in the web genre, the type *tolerate* is preferred over the type *put up with* by American web developers. As illustrated in Table 1, in the web genre, the frequency of *tolerate* is 1,146 tokens, whereas that of *put up with* is 666 tokens.

It is vital that *tolerate* ranks fourth (1,061 tokens) in the academic genre, whereas *put up with* ranks fourth (666 tokens) in the web genre. Again, in rank-four, two types (*tolerate* and *put up with*) do not reveal the same property, thereby exhibiting no similarity. It should be pointed

out, on the other hand, that in the academic genre, the use of *tolerate* is still higher (1,061 tokens) than that of *put up with* (119 tokens). From this it seems clear that in the academic genre, *tolerate* (1,061 tokens) is favored over *put up with* (119 tokens).

It is interesting to observe that *tolerate* and *put up with* rank fifth (946 tokens vs. 570 tokens) in the spoken genre. Quite interestingly, *tolerate* patterns with *put up with* in the spoken genre, thus exhibiting the same characteristic in rank-five. It must be noted, however, that *tolerate* (946 tokens) is preferred over *put up with* (570 tokens) by Americans. It can thus be inferred that Americans prefer using *tolerate* to using *put up with* in daily conversation.

It is worth pointing out that *tolerate* ranks sixth (870 tokens) in the fiction genre, whereas *put up with* ranks sixth (398 tokens) in the magazine genre. Again, two types (*tolerate* and *put up with*) do not reveal the same characteristic (in rank-six), thereby showing no similarity.

It is interesting to point out that *tolerate* and *put up with* rank seventh (816 tokens vs. 364 tokens) in the newspaper genre. Quite interestingly, *tolerate* patterns with *put up with* in the newspaper genre, hence the same property. It should be noted that the use of *tolerate* is even higher (816 tokens) than that of *put up with* (364 tokens) in the newspaper genre. This in turn means that American journalists prefer to use *tolerate* rather than use *put up with*.

It should be pointed out that *tolerate* ranks eighth (712 tokens) in the TV/movie genre, whereas *put up with* ranks eighth (119 tokens) in the academic genre. Again, in rank-eight, *tolerate* does not pattern with *put up with*, thus revealing no similarity. To sum up, two types (*tolerate* and *put up with*) reveal the same characteristic in two genres, whereas they do not in six genres. It can thus be concluded that *tolerate* is 25% the same as *put up with* in the genre analysis.

Now let us turn to Figure 1 and Figure 2. They show the percentage of the frequency of *put up with* and *tolerate* in each genre:

Figure 1 The percentage of the frequency of put up with
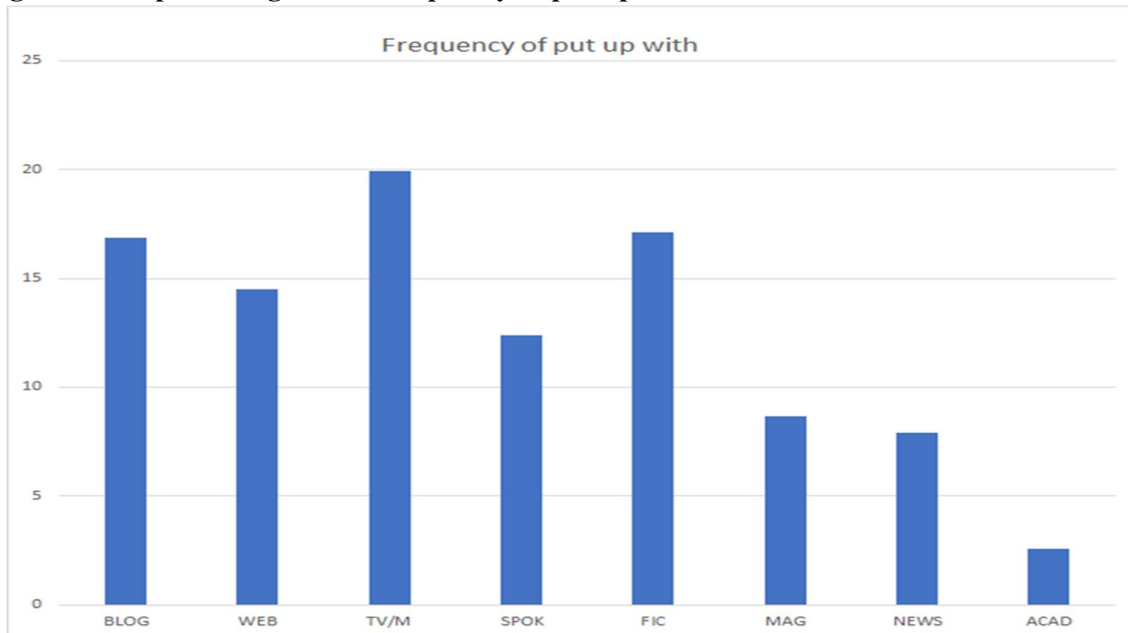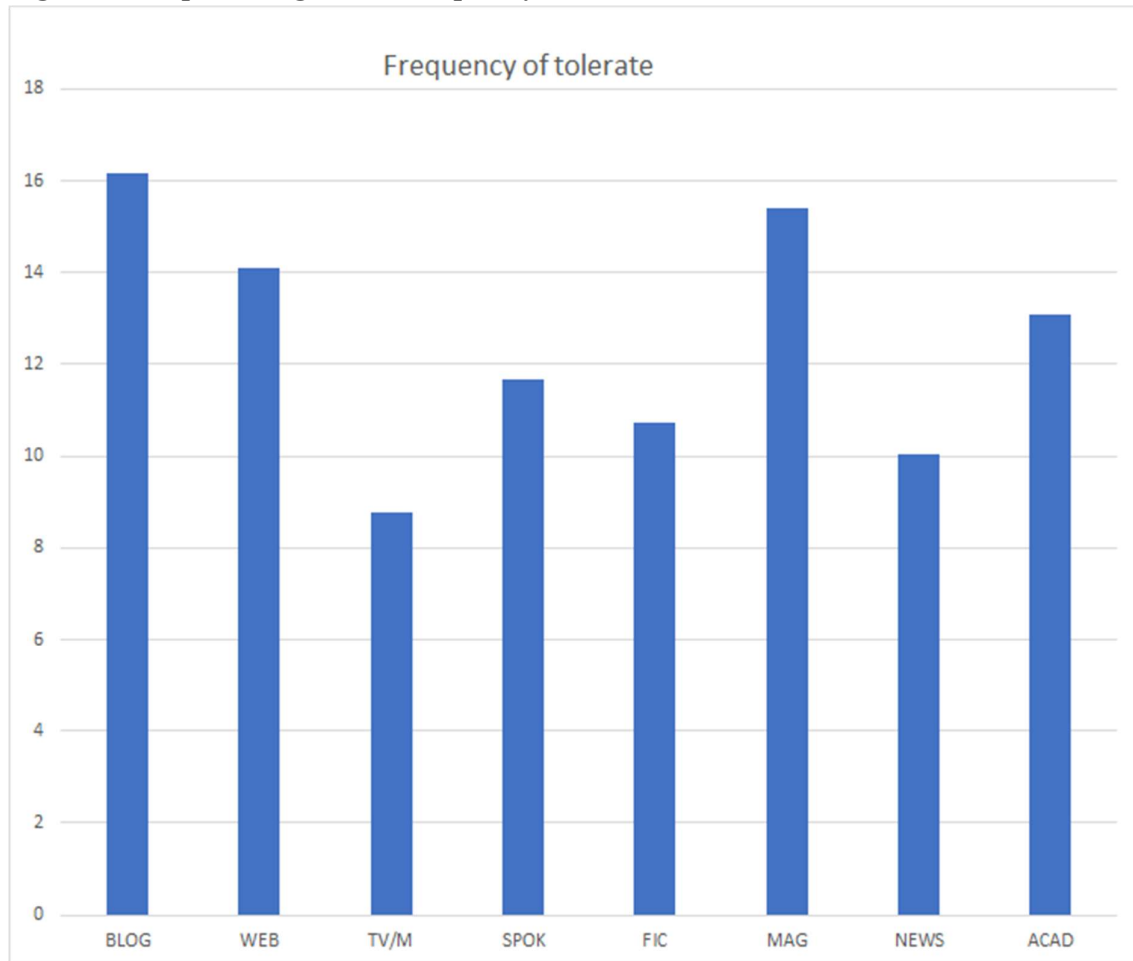

As exemplified in Figure 1, the blog genre was the most influenced by *put up with*, followed by the magazine genre, the web genre, the academic genre, the spoken genre, the fiction genre,

the newspaper genre, and the TV/movie genre, in that order.

Now take a look at Figure 2:

Figure 2 The percentage of the frequency of tolerate

As indicated in Figure 2, the blog genre was the most influenced by *tolerate*, followed by the magazine genre, the web genre, the academic genre, the spoken genre, the fiction genre, the newspaper genre, and the TV/movie genre, in descending order.

2.2. The Euclidean distance

This section is devoted to comparing *tolerate* and *put up with* in terms of the Euclidean distance. By using the Euclidean distance, we can measure the distance between *tolerate* and *put up with* in each genre. The distance between two types in each genre indicates the degree of the similarity between them. As for the Euclidean distance, we adopt the following definition:

(1) Euclidean distance

$$\sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2} = \sqrt{\sum_{i=1}^{n} (p_i - q_i)^2}$$

Now let us observe the distance between *tolerate* and *put up with* in each genre:

Table 2 Euclidean distance between tolerate and put up with

GENRE	BLOG	WEB	TV/M	SPOK	FIC	MAG	NEWS	ACAD
Percentage of tolerate	16.18	14.12	8.77	11.66	10.72	15.39	10.05	13.07
Percentage of put up with	16.88	14.49	19.93	12.4	17.12	8.65	7.91	2.58
Euclidean distance	0.7	0.37	11.16	4.26	6.4	6.74	2.14	10.49

Quite interestingly, the type *tolerate* is the furthest from the type *put up with* in the TV/movie genre. More specifically, in the TV/movie genre, the Euclidean distance between *tolerate* and *put up with* is 11.16, which is the highest among eight genres. It amounts to saying that two types (*tolerate* and *put up with*) have the lowest similarity. More interestingly, the type *tolerate* is the nearest to the type *put up with* in the web genre. To be more specific, in the web genre, the Euclidean distance between *tolerate* and *put up with* is 0.37, which is lowest. From this it seems evident that *tolerate* and *put up with* have the highest similarity in the web genre. As for the Euclidean distance between *tolerate* and *put up with*, the web genre is followed by the blog genre. More specifically, the Euclidean distance between *tolerate* and *put up with* in the blog genre is 0.7, which is the second lowest. It can thus be inferred that *tolerate* and *put up with* have a high similarity in the blog genre. We thus conclude that the web genre has the highest similarity and the blog genre, the newspaper genre, the spoken genre, the fiction genre, the magazine genre, the academic genre, and the TV/movie genre follow.

2.3 The Collocations of tolerate and put up with in the COCA

This section focuses on comparing two types (*tolerate* and *put up with*) in the COCA. We aim to examine the collocation of *tolerate* and that of *put up with* and compare them in the COCA. By doing so, we can capture the similarity between *tolerate* and *put up with* in the COCA:

Table 3 Collocation of tolerate in the COCA

Number	Collocation	Frequency
1	tolerate people	32
2	tolerate dissent	27
3	tolerate plan	23

4	tolerate racism	21
5	tolerate discrimination	20
6	tolerate violence	16
7	tolerate frustration	15
8	tolerate drought	14
9	tolerate failure	14
10	tolerate attacks	13
11	tolerate harassment	13
12	tolerate temperatures	11
13	tolerate behavior	10
14	tolerate abuse	10
15	tolerate mistakes	10
16	tolerate things	10
17	tolerate dairy	9
18	tolerate uncertainty	9
19	tolerate water	9
20	tolerate ambiguity	8
21	tolerate criticism	8
22	tolerate diversity	8
23	tolerate gays	8
24	tolerate name	8
25	tolerate risk	8

Most importantly, *tolerate people* has the highest frequency (32 tokens). This in turn indicates that the expression *tolerate people* is the most preferable one (32 tokens) for Americans. Note that *tolerate peopl*e is followed by *tolerate dissent.* The frequency of the latter is 27 tokens and it ranks second (the second highest). Interestingly, the expression *tolerate dissent* is followed by the expression *tolerate pain*. This in turn implies that *tolerate pain* is the third most frequently used (23 tokens). It is important to note that *tolerate racism* ranks fourth (21 tokens) in the COCA. Quite interestingly, *tolerate discrimination* ranks fifth (20 tokens) in the COCA. We take this as meaning that the expression *tolerate discrimination* is the fifth most widely used. We thus conclude that *tolerate people* is the most preferred (32 tokens) by Americans, followed by *tolerate dissent*, *tolerate pain*, *tolerate racism*, and *tolerate discrimination*, in that order.

Now we aim at comparing the collocation of *tolerate* and that of *put up with* in the COCA:

**Table 4 Collocation of put up with in the COCA**

Number	Collocation	Frequency
1	put up with people	23
2	put up with things	18
3	put up with crap	10
4	put up with abuse	7
5	put up with shit	6

6	put up with stuff	6
7	put up with fools	4
8	put up with me	4
9	put up with hardship	3
10	put up with harassment	3
11	put up with discomfort	3
12	put up with it	3
13	put up with bullshit	3
14	put up with questions	3
15	put up with years	3
16	put up with coyotes	2
17	put up with assholes	2
18	put up with appeasement	2
19	put up with ads	2
20	put up with insults	2
21	put up with friends	2
22	put up with delays	2
23	put up with day	2
24	put up with men	2
25	put up with noise	2

Most interestingly, *put up with people* has the highest frequency (23 tokens). This amounts to saying that the expression *put up with people* is the most preferred (23 tokens) by Americans. Interestingly, *put up with things* ranks second (18 tokens) in the COCA, which we take as indicating that the expression is the second most widely used. It is interesting to note that *put up with crap* ranks third (10 tokens) in the COCA. It is thus appropriate to assume that *put up with people* is the most preferable one (23 tokens) for Americans, followed by *put up with things*, *put up with crap*, and *put up with abuse*, in descending order. It is worthwhile noting that *tolerate things* ranks thirteenth and *put up with things* ranks second. It is worth mentioning, on the other hand, that *tolerate abuse* ranks thirteenth, whereas *put up with abuse* ranks fourth. Finally, it is significant to note that *tolerate harassment* ranks tenth, whereas *put up with harassment* ranks ninth. To sum up, both *tolerate* and *put up with* have the nouns *people*, *things*, *abuse*, and *harassment* as their collocations. That is to say, 4 of 46 nouns are the collocations of both *tolerate* and *put up with*. Note that our list was cut off in the top 25. This in turn indicates that *tolerate* is 8.69% the same as *put up with* in the collocation analysis of the COCA.

Now we attempt to capture this fact in terms of the software package NetMiner. The NetMiner provides the network of the collocations of *tolerate* and *put up with* :

Figure 3 Visualization of the collocations of tolerate and put up with
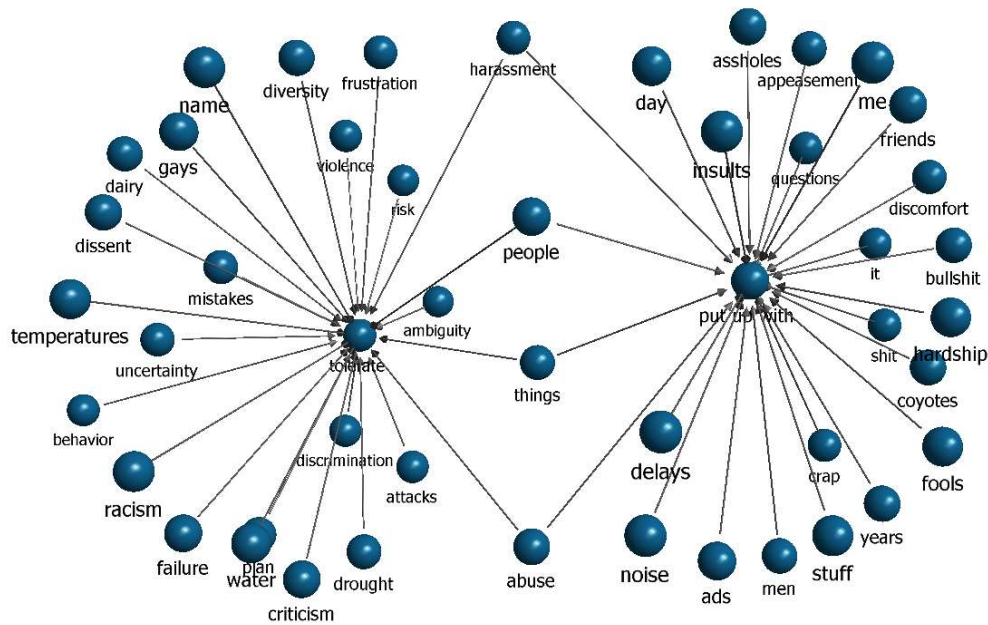
More interestingly, 21 nouns are linked to *tolerate* and *put up with*, respectively. Most importantly, 4 of 46 nouns are linked to both *tolerate* and *put up with*. More specifically, the nouns *people*, *things*, *abuse*, and *harassment* are the collocations of both *tolerate* and *put up with*. For the network of synonyms and big data, see Kang (2022a, 2022b, 2022c, 2022d, 2023a, 2023b). It can thus be concluded that *tolerate* and *put up with* exhibit a very low similarity in the collocation analysis (the top 25) of the COCA.

3. Conclusion

To sum up, we have provided a comparative analysis of *tolerate* and *put up with* in the Corpus of Contemporary American English (COCA). In section 2.1, we have shown that two types (*tolerate* and *put up with*) reveal the same characteristic in two genres, whereas they do not in six genres. More specifically, *tolerate* is 25% the same as *put up with* in the genre analysis. In section 2.2, we have argued that in the TV/movie genre, the Euclidean distance between *tolerate* and *put up with* is 11.16, which is the highest among eight genres. It amounts to saying that two types (*tolerate* and *put up with*) have the lowest similarity. We have further argued that the type *tolerate* is the nearest to the type *put up with* in the web genre. To be more specific, in the web genre, the Euclidean distance between *tolerate* and *put up with* is 0.37, which is lowest. From this it seems evident that *tolerate* and *put up with* have the highest similarity in the web genre. In section 3.3, we have shown that *tolerate people* and *put up people* have the highest frequency (32 tokens vs. 23 tokens). Finally, we have maintained that the nouns *people*, *things*, *abuse*, and *harassment* are the collocations of both *tolerate* and *put up with*. It can thus be inferred that *tolerate* and *put up with* exhibit a very low similarity in the collocation analysis (the top 25) of the COCA.

References

[1] Kang, N. (2022a). A Comparative Analysis of Search for and Look for in Four Corpora. *Advances in Social Sciences Research Journal*, *9*(3), 168-178.

[2] Kang, N. (2022b). A Comparative Analysis of Impressed by and Impressed with in Two Corpora. *Theory and Practice in Language Studies*, *12*(5), 819-827.

[3] Kang, N. (2022c). On Speak to and Talk to: A Corpora-based Analysis. *Theory and Practice in Language Studies*, *12*(7), 1262-1270.

[4] Kang, N. (2022d). On Speak with and Talk with: A Corpora-based Analysis. *International Journal of Social Science and Human Research*, *5*(8), 3354-3360.

[5] Kang, N. (2023a). K-Pop in BBC News: A Big Data Analysis. *Advances in Social Sciences Research Journal*, *10*(2), 156-169.

[6] Kang, N. (2023b). K-Dramas in Google: A NetMiner Analysis. *Transaction on Engineering and Computing Sciences*, *11*(1), 193-216.