Vol. 4 | No. 4 | Page 629 – 647 | 2025 |

https://journals.su.edu.ye/index.php/jhs

ISSN: 2958-8677

### The Investigation of the Frequency and Genre Distribution of the English Synonymous Verbs Denoting Change in Light of the Corpus of Contemporary American English

### Ali Ahmed Hizam Julul <sup>1,\*</sup>, Yahya Mohammed Ali Al-Marrani <sup>2</sup>

### **Keywords**

- 1. COCA
- 3. Frequency
- 5. Near-synonym

- 2. Corpus Linguistics
- 4. Genre

### **Abstract:**

Preferring one synonym to another is a daunting challenge for EFL learners. This corpus-based study aims at analyzing the frequency and genres of the English synonymous Verbs denoting change utilizing the corpus-based as well as the qualitative and quantitative methodology. The findings demonstrate that the target verbs in this study differed from one another in terms of their overall frequency in the Corpus of Contemporary American English (COCA) to a great extent. From highest to lowest, the target verbs are ranked in COCA in this order, change, adjust, adapt, alter, and modify, respectively. Of the genre distribution, aside from change which tend to emerge in both formal and informal discourses, the other verbs have a propensity to appear most commonly in formal discourses and thus they are marked with high degree of formality. Additionally, formality of these verbs is confirmed by Coxhead's (2000) except for adapt which is absent from Coxhead's academic words though reported in this study to be most frequent in the academic texts of COCA. The specificities of these verbs enhance the claim that near-synonymous verbs cannot be used interchangeably in all contexts. Hence, course designers and teachers are recommended to utilize corpora to teach synonyms considering their frequency and genres as this information is overlooked by contemporary dictionaries.

<sup>&</sup>lt;sup>1</sup>Department of English-Faculty of Languages - Sana'a University, Sana'a, Yemen.

<sup>&</sup>lt;sup>2</sup>Department of English-Faculty of Education-Arhab - Sana'a University, Sana'a, Yemen.

<sup>\*</sup>Corresponding author: jululmind@gmail.com & almarrani99@gmail.com

https://journals.su.edu.ye/index.php/jhs

**ISSN: 2958-8677** 

# استكشاف التردد والأنواع الأدبية للأفعال المترادفة الدالة على التغيير في اللغة الإنجليزية في ضوء المدونة اللغوية للانجليزية الأمريكية المعاصرة

### على أحمد حزام جعلل 1 \* , يحيى محمد على المرانى 2

اقسم اللغة الانجليزية ، كلية اللغات - جامعة صنعاء ، صنعاء ، اليمن. 2قسم اللغة الانجليزية، كلية التربية ارجب - جامعة صنعاء ، صنعاء ، اليمن.

\*المؤلف: Jululmind@gmail.com في jululmind@gmail.com

#### الكلمات المفتاحية

كوكا
لسانيات المدونات اللغوية
الأنواع الأدبية

5. مرادف جزئي

### الملخص:

يُعدُ تفضيل مرادف على آخر تحديًا صعبًا لمتعلمي اللغة الإنجليزية؛ لذا تهدف هذه الدراسة إلى تحليل تردد الأفعال المترادفة الدالة على التغيير في اللغة الإنجليزية والأنواع الأدبية التي ترد فيها. وقد تم اتباع منهج المدونات اللغوية بالإضافة الي المنهج الكمي والنوعي في تحليل البيانات، تُظهر النتائج أن الأفعال المستهدفة في هذه الدراسة تختلف عن بعضها بعضاً من حيث التردد الكلي إلى حد كبير، فقد اظهرت النتائج ترتيب هذه الأفعال من الأعلى إلى الأدنى بناء على ترددها الكلي في المدونة الأمريكية المعاصرة للغة الإنجليزية (كوكا) كما يأتي، يغير ( change)، يغير ( change)، يغير ( (adapt)، يعدل ((digla) للمستهدفة النافع المستهدفة المدونة اللغور بصورة أكثر شيوعًا في الأنواع الأدبية الرسمية وبالتالي فهي تتميز بدرجة عالية من الرسمية. بخلاف ذالك أظهرت النتائج في هذه الدراسة مع قائمة كوكسهيد ((2000) للكامات الرسمية حيث ظهرت هذة الأفعال في هذة القائمة بالإضافة إلى ذلك، تتوافق النتائج في هذه الدراسة أن الأفعال المستهدفة من حيث الأنواع الأدبية دليلاً على أن الأفعال المستهدفة من حيث التردد والتوزيع في الأنواع الأدبية دليلاً على أن الأفعال المستهدفة من حيث التردد والتوزيع في الأنواع الأدبية دليلاً على أن الأفعال المترادفة لا يمكن استخدامها بالتبادل في جميع المياقات اللغوية لذا فإنَّ أغلب الترادف في اللغة الإنجليزية يكون جزئيًا بين الكلمات. ختاماً أوصت الدراسة مصممي المقرات الدراسية والمعلمين باستخدام المدونات اللغوية في تدريس المترادفات بناءً على تكرارها وأنواعها الأدبية حيث أن هذه الدراسة الميشار اليها في القواميس المعاصرة.

مجلة جامعة صنعاء للعلوم الإنسانية

#### **Introduction:**

To learn, teach or even translate English language effectively, vocabulary must be given substantial attention. The more vocabulary the learners and teachers master, the more effectively and proficiently they can express themselves (Yusu, 2014) and (Szudarski, 2018). Put simply, acquiring a great deal of vocabulary is essential for effective communication in English language (Sayyed & Al-Khanji, 2019). In a similar vein, Phoocharoensil (2020) confirmed that vocabulary learning is essential to the second language acquisition.

One aspect of vocabulary difficulty encountered by the EFL learners is the underlying linguistic differences among the near synonymous words. The recognition of such underlying aspects might help to avoid causing the listeners and readers to misunderstand the English discourse. According to Sormet (2017) and Song (2021), English language has a large stock of synonymous vocabulary. One source is the borrowing of a lot of words from different languages such as Arabic, French, Greek and Latin. However, English synonyms are not interchangeable in all contexts. Thus, EFL learners should be aware of the fact that English synonyms, though they have approximately the same meaning, are used differently in several linguistic contexts and aspects such as their frequency and distribution in different genres (Kiatthanakul, 2015).

In recent times, there has been a consensus among several scholars such as Aroonmanakun (2015), Sridhanyarat (2018), Phoocharoensil & Kanokpermpoon,(2021) and Chaokongjakra (2023) that learners of English, more specifically second and foreign language learners, get confused to select which synonym should be suitably used in particular context. To put it differently, second and foreign language learners of English, even those at the advanced language proficiency level, are muddled about the choice of synonymous words, as they fail to recognize the slight differences among them (Lee & Liu,

2009). This might be attributed to the fact that second language learners misconstrue synonyms being interchangeable in all contexts (Phoocharoensil & Kanokpermpoon, 2021). In addition to this, traditionally, dictionaries are the main sources for language teachers and learners to differentiate synonyms. However, there is an absence of detailed information of nearsynonyms such as their distribution among genres, collocations and colligations. Generally, dictionaries tend to offer the core meanings of the concepts of synonyms with few examples (Song, 2021). A considerable number of scholars such as Lee & Liu (2009), Ly and Jung (2015) and Lertcharoenwanich (2023) are in keeping with this claim in that dictionaries present synonyms in a way that implies their interchangeable status. Though a reference to some collocations, and colligations is provided in dictionaries, the frequency and genre distribution of the words are absent or just obliquely connoted. Thus, the major aim of this study is to further demonstrate the efficacy of corpora, more specifically, COCA in unveiling the similarities and specificities among the near synonyms in terms of frequency and genre distribution.

The advent of corpus linguistics has led to a great shift in the study of vocabulary. Language pedagogues and researchers can use a corpus to conduct an in-depth scrutiny of language vocabulary, more specifically synonyms based on their frequency and genre distribution. Employing corpora to investigate the similarities and differences among the synonyms based on frequency and genre is regarded as the most appropriate approach due to the fact that it has been proved effective, reliable, and empirical by several scholars such as Yuliawati & Indira (2019), Li (2019); O'Keeffe & McCarthyv (2010). Interestingly, a growing body of literature such as Jarunwaraphan & Mallikamas ( 2020), Song (2021), Sumonsriworakun (2022), Lertcharoenwanich (2023) and Chaokongjakra (2023) has conclusively established that the corpus-based analysis of synonyms based on

their frequency and genre distribution is practical enough to reasonably and evidently distinguish the near synonyms from one another.

In view of all that has been mentioned so far, and since there has not been a corpus-based study yet on the differences among the English synonymous verbs denoting change, this study aims to unveil such similarities and differences in terms of frequency and genre distribution. And for this objective it seeks to answer the following questions:

- 1. What are the similarities among the English synonymous verbs denoting change in terms frequency and genre distribution?
- 2. What are the differences among the English synonymous verbs denoting change in terms frequency and genre distribution?

### **Literature Review**

### **Synonyms**

"Synonyms are words or expressions that have the same meaning in some or all contexts" (O'Grady & Archibald, 2016, p. 182). For Halliday, Teubert, Yallop, & Cermakova (2004, p. 170), "synonymy is a relationship of near identity in meaning as in dentures and false teeth". Several scholars such as Jarunwaraphan & Mallikamas (2020), Phoocharoensil (2022), Chaokongjakra (2023) and Lertcharoenwanich (2023) have established that synonyms are categorized into two types, namely perfect and near synonyms. Perfect synonyms are words that have approximately the same meanings and can be replaced in the same contexts (Pijuntug, 2015) . However, near-synonyms are "lexical items whose senses are identical in respect of central semantic traits, but differ in minor or peripheral traits" (Chung, 2011, p.400). For Phoocharoensil (2020), near-synonyms are two words that overlap in some meanings but cannot substitute one another in all contexts as each synonym has its own linguistic specificities such as frequency, genre, collocation and colligation. Several recent studies have confirmed that English language has no prefect synonymous words that can be applied interchangeably in all contexts (AlAmro, 2019), (Ajmal, Kumar, Ritonga, & Nukapangu, 2022), (Boontam & Phoocharoensil, 2022), Kruawonga & Phoocharoensil, 2022) and (İŞLER, 2022). O'Grady & Archibald (2016, p.182) stated "Because it would be inefficient for a language to have two words or phrases with absolutely identical meanings, perfect synonymy is rare, if not impossible". Mikhailov & Cooper (2016) also pointed out that these is an absence of perfect synonyms in English as each synonym contains its own unique shade of meaning. Thus, roughly all the synonyms in English language are characterized as near synonyms that cannot be used interchangeably in all contexts.

## Corpus linguistics and English synonyms

A large and growing body of literature on corpus linguistics such as Baker, Hardie, & Mcenery (2006), Dash (2008), Biber & Randi (2015), and Hunston (2022) has described corpus linguistics as a new approach of research through which language use and variations can be investigated empirically yielding research results of greater validity and generalizability. Sinclair (1991, p. 100) suggested that by gathering a lot of data and exploring it at once, corpus linguistics provides insightful perspectives on everyday discourse and enhance the view that variations in language is systematic that could be explained through empirical, and quantitative approaches. Corpus linguistics makes it possible for researches to investigate the frequency of synonymous words or phrases, in different genres of texts such as academic, fiction, spoken language, and journals and thus reach reasonable and evident conclusions of similarities and specificities among synonymous words (Jarunwaraphan & Mallikamas, 2020).

### The Criteria to Distinguish Near Synonyms

### Frequency

The concept of frequency in corpus linguistics refers to analyzing the words by exploring their number of occurrences within corpora. Frequency analyses also allow researchers and linguists to make comparisons between different words in a corpus discovering, for instance, the most or least common words in written or spoken English, or how has particular words changed over time. Furthermore, linguists have made keyword lists, composed by frequency counts of words in corpora (Baker, Hardie, & Mcenery, 2006).

Through corpora, most frequent words that should be taught first are listed to facilitate language learning. The frequency practicality and lexical relationship have empirically made identification of the most common lexical items much easier and more data-informed than before (Martinez & Schmitt, 2015). However, when we compare the frequencies of words in different or the same corpora we could reach qualitative results that make it possible to realize the fact that particular words are more frequently used than others. Corpus linguistics statistic tools and analysis programs make it possible researchers to easily explore the frequency of the words and thus determine which ones are highly frequent in use and should gain preference over other less frequent ones in learning and teaching of a language. The frequency analysis is also of a great importance to language learners in that it helps in categorizing a word into high, mid or low frequency and thus leading the learners to throw the focus on the core words that form a great percentage of a language needed in the daily communication (Szudarski, 2018).

#### Genre

The category of a text is the most essential organizing norm in the most contemporary

corpora. Even early corpora, such as the Brown, and London-Lund corpora, are structured considering the text types. For example, the Brown corpus is organized to signify text categories like press editorials, press reportage, biographies, essays, academic prose, common fiction (Biber, 2010). The term genre comes from Latin which means class. In linguistics, the word signifies distinctive type of text. Corpora are constructed of various genres such as fiction, religious, press, academic, and private letters. A genre may be sub-classified into a sub-genre, for instance, fiction may be subcategorized into science fiction, mystery, humor and westerns romance (Baker, Hardie, & Mcenery, 2006). Cheng (2012) explained that genres are culturally acknowledged forms of language associated with particular situations such as phone calls and research articles. The genre has been discussed by Biber & Conrad (2009) who explained that the concept refers to text categorizing that is defined situationally. With regard to the genres included in the corpus on which this study is based, the 2020 updated version of COCA encompasses over a billion words in 485,202 texts distributed evenly in the following eight genres, blogs, web, TV and movie, spoken, fiction, popular magazines, newspapers, and academic texts. The balance across the eight genres from year to year insures providing data that is highly reliable and is relatively different from data provided by other genre-balanced corpora that are not (Phoocharoensil S., 2021).

### Previous Corpus-Based Studies on English Synonyms

Several studies such as Sumonsriworakun (2022), Chaengchenkit (2023) and Sridhanyarat & Phoocharoensil (2023) pointed out that the genres differ in the degree of formality. While TV and Movie Subtitles genre and Spoken genre contain very informal or even colloquial texts, Academic genre comprises formal texts. The Magazines and Newspapers genre consists of

different topics such as international, national and local news, opinion, sport and so forth. Thus, it contains both formal and informal texts. Accordingly, most of the previous corpus-based studies on synonyms differentiated among synonyms based on their degree of formality considering the proportion of their distribution in corpus genres.

In recent years, there has been a considerable number of corpus-based studies that have been conducted on near-synonyms with a view to unveiling their similarities and differences based on their frequency and genre. To begin with, investigated Phoocharoensil (2020)similarities and differences among the nearsynonyms: consequence, result, and outcome based on their distribution across eight genres in COCA. The findings highlighted that these near synonyms are used with the highest frequency in academic texts, which is evident that they are all associated with a high degree of formality. More evidence is seen in their low frequency in informal genres such as TV and movie subtitles. In line with this Kruawonga & Phoocharoensil (2022) concluded that the near synonyms teach, educate and instruct have a propensity to emerge in formal genres, more specifically academic texts. This observation is further supported by the low number of occurrences in informal genres, such as spoken, fictions and TV/Movie. With respect to the proportion in frequency, teach has a tendency to appear much more extensively and generally than educate and instruct across the eight genres of COCA. The findings of Phoocharoensil (2022) is also in keeping with this observation in that the near synonyms primary, main, and major tend to abound in academic texts that are characteristic of a very high level of formality. The relative infrequency of the three near synonyms in fiction and TV/movie subtitles confirms that these words are associated with formal texts. Moreover. Sridhanyarat & Phoocharoensil (2023) reported that assess, evaluate, and measure are most frequently used in academic texts. Such evidence enhances the claim that these near synonyms are marked with formality. Additionally, their formality is confirmed by the fact they are produced least frequently in less formal settings such as spoken, fiction, and TV and movie subtitles.

However, a number of studies highlighted how near synonyms differ significantly from one another in terms of frequency and genre distribution in corpora. Jarunwaraphan & Mallikamas (2020), for example, clarified how opportunity is preferred to chance in academic texts seeing that it scored a higher number of occurrences in COCA. Moreover, the high frequency of chance in informal contexts such as spoken language and fiction seem to imply the difference in the degree of formality between the two near synonymous nouns. In keeping with this, Sumonsriworakun (2022) explained that disadvantage predominantly emerges in formal or academic texts, whereas downside drawback have a tendency to appear in informal Lertcharoenwanich (2023) discourse. concluded that the synonymous adjectives: blank, empty, and vacant behaved differently in terms of frequency and genre distribution. With respect to the frequency, empty tends to occur with the highest frequency, followed by blank and vacant, respectively. Of the genre distribution. blank and empty tend to emerge most frequently in fiction while vacant has a propensity to appear in newspapers. In addition to this, these synonymous adjectives are uncommon in formal discourse such as academic texts. Due to the specificity of each adjective, they cannot be used interchangeably in all contexts. Given that the previous studies have proved the efficiency of corpus linguistics as a method in differentiating the near synonyms based on their frequency and genre distribution, the current study aims to uncover the similarities and differences among the English verbs denoting change, namely change, alter, adapt, adjust and modify in light of the Corpus of Contemporary American English (COCA).

### Methodology

This study has applied a corpus linguistic approach as a methodology to measure the similarities and differences among the target verbs in terms of frequency and genres. Tognini Bonelli (2001) clarified that, in a corpus-based study, corpus linguistics is a methodology or a practical tool to discover several aspects of the natural use of language that might go unnoticed. Mccarthy, Carter (2007)O'keeffe. & demonstrated that both qualitative quantitative analyses are applied in corpus studies. For instance, when exploring the word number of occurrences, quantitative results are supposed to appear. However, when the frequencies of the same words are compared in different corpora or the frequency of different words in the same corpora, qualitative results are yielded.

In this study, the Oxford Advanced Learner's Dictionary 10<sup>th</sup> edition and Longman Dictionary of Contemporary English 6th edition are used as the primary sources for the basic information of the target words. In addition to this, the updated version of the Corpus of Contemporary American English (COCA, 2020) is selected as a source of data due to the following substantial reasons. According to Davies (2010) and Thongpan (2022), COCA is recognized as one of the best, resourceful and most-widely used corpora in authentic American English for the purpose of English language learning, teaching, research and practice. The 2020 updated version of COCA encompasses over a billion words distributed evenly in the following eight genres, blogs, web, TV and movie, spoken, fiction, popular magazines, newspapers, and academic texts. The balance across the eight genres from year to year insures providing data that is highly reliable and is relatively different from data provided by other corpora that are not genrebalanced (Phoocharoensil, 2021). Moreover, due to the constant update of COCA by adding new texts annually, it is recognized as a monitor corpus which monitors the progress of American English diachronically (Kayaoglu, 2013) (Phoocharoensil S., 2020). Additionally, COCA has a free, online, web-based interface. Thus, it can be easily accessed by anybody who has an access to the internet (Chaengchenkit, 2023).

In this study, an in-depth scrutiny of the English verbs denoting change was conducted in accordance with their overall frequency and distribution in the COCA genres. As explained this study aimed at unveiling how the target synonyms: change, alter, adapt, adjust, and modify behaved across the eight genes in COCA, with the light shed on formality and informality of the genres in which the target synonyms tend to abound. Given that the verbs have different inflectional forms, they were lemmatized by inserting the target verb in capital letters in the search box of COCA interface so that the overall frequency of each verb lemma is considered in the results produced by COCA. To clarify, by inserting CHANGE in COCA search box, the inflectional forms of the verb change such as change, changes, changing and changed are all counted in COCA automatic search.

### **Results and Discussion**

### The Overall Frequency of the Target Verbs

To thoroughly understand the similarities and differences of English verbs denoting change based on COCA, the first and most primary information is indubitably their overall frequency. It is of note that since the verbs have different inflectional forms. thev lemmatized so that the overall frequency of each verb lemma is considered in the results as shown in table 1. In addition to this, due to the fact that COCA is a mega corpus containing roughly a billion words, it is more observable to show how many times each word occurs in a million besides the total frequency figures. It is also worth noting that both the total frequency as well as the normalized frequency per million of any word is provided in the charts of COCA. The following

equation elucidates how the frequency of each word is normalized per million.

the frequency of a word per milliom  $= \frac{\text{the total frequency of the word x 1000000}}{\text{the total corpus size (993000000)}}$ 

Thus, the data reported in table1 illustrates that the ranking of the target verbs frequency in COCA from lowest to highest respectively as follows *modify* has a total frequency of 15996 times with an average of 16.11 times per million words, *alter* has a total frequency of 24907 times with an average of 25.08 times per million words, *adapt* has a total frequency of 24934 times with an average of 25.11 times per million words, *adjust* has a total frequency of 33203 times with an average of 33.44 times per million words, whereas *change* has total occurrences of 327685 times with an average of 329.99 times per million words.

Table 1: The total frequency of the target verbs in COCA

No	Verb	Total Frequency	Corpus Size	Per Million	
1	Modify	15996		16.11	
2	Alter	24907	993 Million	25.08	
3	Adapt	24934	773 Willion	25.11	
4	Adjust	33203		33.44	
5	Change	327685		329.99	

An analytical glance at table 1 discloses that the five verbs, though they all signify the act of change, vary sizably in their overall frequency. Surprisingly, change appears with the highest frequency among the five verbs scoring 329.99 times per million while modify occurs with the lowest frequency scoring only 16.11 times per million words. Both alter and adapt happen to score approximately the same in-between ranks occurring 25.08 and 25.11 times per million respectively. Additionally, adjust scored a midposition occurring 33.44 times per million words which is to some extent higher than alter and adapt. Another observation is that change having the greatest frequency among the five verbs (329.99 times per million) is almost twenty timeshigher than that of *modify* (16.11) which is the lowest in frequency and nearly ten times higher than that of adjust (33.44) which followed change in frequency. The highly great frequency of change may not be surprising given that it denotes various senses and has a broad range of uses. It is worth noting that the broad frequency

of change may imply its wider collocational and colligational range and a more diverse semantic prosody and preference profile. In case of the remaining verbs, it remains questionable how they vary greatly in frequency though they all denote the act of change and have roughly a similar range of denotations. To illustrate, according to Oxford advanced Learner Dictionary 10th edition (OALD 10th ed) and Longman Dictionary of Contemporary English 6th edition (LDOCE 6th ed), adjust, adapt, alter and modify overlap in the sense that a change is made to something so that it becomes suitable for a different purpose, use or situation. The great and significant variations in the frequency of these verbs might imply significant differences in their genre, collocations, colligations, and semantic prosody and preference.

According to the OALD 10<sup>th</sup> ed, the 3000 core words have been selected based on their frequency in the language and their essential relevance to learners of English. These words have been assigned gradual levels in which under

A1 and A2 the elementary and pre-intermediate words are listed. Under B1 and B2, lower-andupper intermediate words are listed. Concerning the target verbs in this study, change is found to be the most frequent verb as it occurs in the elementary list A1. Adapt, modify and alter come in the lower and upper-intermediate list while adjust occurs in a little bit higher level above the 3000 most frequent essential Additionally, In the LDOCE 6th ed, the verb change is listed in the top 1000 spoken and written words. Alter is listed in the top 3000 spoken and written words. Adjust, and adapt occur in the top 3000 most frequent written words while *modify* is listed in a little bit higher level above 3000. From the discussion above it becomes obvious that both dictionaries confirm that the verbs denoting change in the English language are very essential for learners of English, especially EFL learners. Hence, these verbs should be learnt in the primary levels as they occur most frequently in primary English discourse and provide insightful knowledge on the unique patterns and meanings in naturallyoccurring language.

Interestingly, with respect to the overall frequency of the target verbs, the findings of this study as already apparent in table approximately match those in LDOCE 6th ed and OALD 10<sup>th</sup> ed in that all these verbs are used with high frequency in English discourse. However, based on their highest frequency in COCA, the target verbs are ranked in this order, change, adjust, adapt, alter, and modify, respectively. In contrast, the frequency ranking of these verbs emerges in OALD 10<sup>th</sup> ed as *change*, *adapt*, *alter*, modify and adjust. It is of note that in OALD 10<sup>th</sup> ed adjust is ranked the least frequent among the five verbs as it occurs in the upper intermediate level among the 5000 most frequent words and modify is ranked among the most frequent written words within the 3000 level. In the LDOCE 6<sup>th</sup> ed the target verbs appear in this order: *change*, adjust, adapt alter and modify. It may be the case therefore that this inconsistency is due to the fact that the two dictionaries were built based on two different corpora that encompass American, British and world Englishes while the findings of this study is only based on COCA which is pure representation of the American English.

The findings of this study are in agreement with those obtained by Jones & Waller (2015) which highlight that the frequency analysis of words based on corpora data has been proved advantageous in teaching and learning English as it shed the light on the words that should be given preference and priority and thus taught first. Applying the intuition of teachers and learners has failed in prioritizing words for learning. Naturally, after investigating the overall frequency of the target verbs in this study, it could be suggested that the verb change must be learnt first being the most frequent followed by adjust, adapt, alter and finally modify. Although. such analysis has been provided by both dictionaries employed in this study and confirmed by the corpus findings in this study, other aspects such as genre distribution, collocations, colligations, semantic preferences and semantic prosody can only be uncovered through corpus analysis. This enhances the claim that profound information is obtained the time corpus-based studies are applied to explore near synonyms (Hunston, 2002).

To briefly recapitulate this section, the previous lines has revealed that the target verbs in this study differed from one another in terms of their overall frequency of occurrences in COCA to a great extent which enhances the claim of many scholars such as Phoocharoensil (2022), and Ajmal, Kumar, Ritonga, & Nukapangu (2022). that absolute synonyms scarcely occur in English. Another observation is that the findings of this study are in line with the data in both OALD 10<sup>th</sup> ed and LDOCE 6<sup>th</sup> ed with a slight inconsistency. The principal limitation of this analysis was that the total frequencies of the target verbs were retrieved solely from COCA which is only representative of American English. Future studies might include other corpora such as BNC which is representative of British English. It is supposed that the findings may achieve reasoning behind the questionable aspects that remained unclear or yield further findings of considerable significance.

### The Genre Distribution of the Target Verbs

In this section, the analysis, the arguments, and findings meet the part of the research questions that concern the similarities and differences among the target verbs based on genre distribution. Chaengchenkit (2023) and Davies (2010) pointed out that the genres differ in the degree of formality. While TV and Movie Subtitles genre and Spoken genre contain very informal or even colloquial texts, Academic genre comprises formal texts. The Magazines

and Newspapers genres consist of different topics such as international, national and local news, opinion, sport and so forth. Thus, they contain both formal and informal texts. Accordingly, the target verbs in this study will be categorized as formal or informal based on their distribution among the eight genres. The target verbs that occur more frequently in TV and Movie Subtitles and Spoken genre will be labeled as informal, however the ones that occur more frequently in the Academic genre will be labelled as formal. In the case of the target verbs in this study that occur with high frequency in Magazines Newspapers, Web, Blog, and Fiction genres, deeper scrutiny of the subgenres in COCA or exploring their behavior in the COCA concordance lines should be considered to determine their formality.

Table 2: The target verbs distribution in the COCA genres									
	1st	2nd	3rd	4th	5th	6th	7th	8th	all
Modify	ACAD 47.6	WEB 28.15	MAG 16.38	BLOG 16.13	NEWS 8.38	SPOK 4.94	FIC 4.55	TV/M 3.71	16.11
Mc	ACAD	WEB	MAG	BLOG	NEWS	FIC	TV/M	SPOK	25.08
Alter	55.68	33.07	32.98	23.23	20.49	15.13	10.87	10.35	
Adapt	ACAD 62.5	MAG 36.32	WEB 28.7	NEWS 26.03	BLOG 24.9	FIC 8.73	SPOK 8.44	TV/M 6.5	25.11
Adjust	ACAD 53.4	MAG 48.79	FIC 40.81	WEB 36.72	BLOG 32.98	NEWS 32.57	SOPK 12.15	TV/M 11.9	33.44
Change	WEB 400.74	BLOG 399.42	SPOK 379.55	MAG 310.72	TV/M 308.59	NEWS 300.15	ACAD 272.02	FIC 260.46	330

Table 2 displays the distribution of the target verbs among the eight genres in COCA. The eight genres have been ranked in order to exhibit the frequency of each target verb from highest to lowest. Considering the top four genres in table 2 in which the target verbs are ranked based on their highest frequency, interestingly, apart from change, all the target verbs have the highest

presence in written English occurring most frequently in Academic, Magazines, Web, and Blog genres. Additionally, in the other end, the considerably low frequency of the target verbs is simply noticed in Spoken genres confirming that the target verbs are marked with high degree of formality. To clarify, with the frequency of the target verbs in the Academic genre in mind, table 2 displays that *adapt* is the most frequent verb scoring 62.5 per million. Alter comes the second scoring 55.68 per million followed by adjust which scored 53.4 per million. Modify comes in the fourth rank scoring 47.6 per million. Based on these findings, it could be claimed that the degree of formality of these verbs is construed through their order in which adapt is the most formal followed by alter, adjust and modify.

This observation is confirmed by the lowest frequency of these verbs in the informal contexts, namely Spoken and TV and Movie which are representative of informal or colloquial English.

In particular, *modify* occurs with the lowest frequency scoring 4.94 per million in Spoken, and 3.71 per million in TV and Movie. In a similar vein, the frequency of *alter* is low in TV and Movie (10.87 per million) and Spoken (10.35)

639

per million). Similarly, adapt has low frequency in Spoken (8.44 per million), and TV and Movie In the same line, adjust (6.5 per million). occurs in Spoken (12.15 per million), and TV and Movie (11.9 per million). According to Niwesworakarn, Phoocharoensil, Petkaew, & Intasingh (2023), words can be more or less synonymous with one another considering the overlapping aspects they happen to share. In parallel to this notion, it could be argued that adjust, adapt, modify and alter are claimed to be more synonymous as they all occur with high frequency in the written genres. however, is less synonymous in contrast with the other target verbs given that its frequency is distributed among all the genres differently than the other verbs.

If one takes into account the fact that all the target verbs except for change scored the highest frequency in the academic genre, a deeper investigation of how the target verbs are distributed in the academic subgenres might highlight the similarities and differences among these verbs more deeply. It is of note that the Academic subgenres of the target verbs have been summarized and ranked according to frequency from highest to lowest in table 3.

Table 3: the academic subgenres of the target verbs

Alter		Adjust		Adapt		Modify	
Sci/Tech	78.26	Business	155.06	Education	93.69	Sci/Tech	78.32
Medicine	66.61	Medicine	135.99	Sci/Tech	87.94	Medicine	66.15
Law/PolSci	59.99	Education	67.32	Geog/SocSci	63.76	Education	56.42
Geog/SocSci	53.77	Sci/Tech	63.02	Humanities	58.08	Law/PolSci	49.24
History	53.08	Geog/SocSci	55.57	History	52.03	Geog/SocSci	47.68
Humanities	49.68	Law/PolSci	32.88	Misc	51.48	Business	41.52
Phil/Rel	46.29	History	32.25	Medicine	48.20	Humanities	29.25
Misc	44.63	Humanities	22.53	Phil/Rel	42.85	Phil/Rel	27.67

Business	40.67	Misc	22.42	Business	40.67	History	24.11
Education	39.24	Phil/Rel	19.00	Law/PolSci	31.50	Misc	21.79

Abbreviation:

Sci= science, tech = technology, PolSci = politics science, Phil = philosophy, Misc= miscellaneous ,Geog= geography, SocSci= social science

Already apparent in table 3 is that the frequency of the target verbs fluctuates among the ten academic subgenres. However, notably, overlapping areas appear to strongly enhance the claim that the target verbs are near synonyms. To exemplify, it appears in table 3 that Americans are apt to use these verbs mostly in science and technology contexts. Table 3 shows that the highest frequency of both alter and modify is in the science and technology contexts with 78.26 and 78.32 per million, respectively. Adapt is also used most commonly in the science and technology field scoring the second rank with 87.94 per million. The use of *adjust* in the science and technology field is also noticeable in that it scored the fourth rank with 63.02 per million. Another similarity among the target verbs is in the distribution of these verbs among the medicine field in which alter, adjust and modify scored the second rank of highest frequency with 66.61, 135.99, 66.15 per million respectively. Inconsistently, in this field, adapt scores the seventh rank scoring 48.20 per million.

To conclude, the target verbs, namely *adapt*, *adjust*, *alter and modify*, though, generally, occur with the highest frequency in the academic genre, are distributed unevenly among the academic subgenres in COCA with some overlapping areas. Interestingly, these findings mirror those of the previous studies such as Chaokongjakra (2023), Sridhanyarat & Phoocharoensil (2023), and Phoocharoensil S. (2020) who stated that some synonyms happen to be favored in formal genres, given that their highest frequency is represented in academic texts rather than spoken discourses. This is further reinforced by the low frequency of occurrences in informal genres such as fiction and TV/movie subtitles. However, the

findings in this study are not in keeping with those of Phoocharoensil (2020), Cai (2012) and Sriwangrach (2024) who concluded that some near synonyms emerge in different genres with different degree of formality.

Turning, now, to the genre distribution of change, unlike adapt, alter, adjust and modify, change frequency behaved differently in that it scored the highest rank in Web, followed by Blog, Spoken, Magazine, TV and Movie, News, Academic and Fiction. Given that *change* occurred with high frequency in both Spoken and TV and Movie genres which come in the third and fifth ranks, respectively, *change* is claimed to be informal. This might be enhanced by the fact that Academic genre comes in the seventh rank. However, a quick look at the distribution of the frequency of change among the eight genres shows that unlike the other target verbs in this study whose high frequency noticeably centered in written English and low frequency centered in English, change occurred spoken considerably high frequency in approximately all genres including both that encompass formal and informal texts. To exemplify, change occurred 272.02 per million in academic genre which comes in the seventh rank and 260.46 per million in Fiction which is the last rank. Hence, it might be argued that the informality of change is still not adequately confirmed and might be further investigated in other corpora such as BNC or explored deeply in concordance lines of the target verbs. It is perhaps more accurate to suggest that the idiomatic meanings and broad range of use that change conveys as clarified in both OALD 10th ed and LDOCE 6th ed may justify its high frequency in Blog Spoken, Magazine, and TV and Movie genres. Also noticeable is the fact that

change is listed in the top 1000 spoken and written words in LDOCE 6th ed which can be another justification for its high frequency in approximately all the genres. Notably, the finding that approximately all the target verbs are associated with academic language is consistent with Coxhead's (2000) Academic Word List (AWL) in which modify, alter, and adjust are listed among the academic words. The absence of change in Coxhead's academic word list is also in line with the findings in this study. Surprisingly, adapt which is absent from Coxhead's academic word list is reported in this study to be most frequent in the academic texts. This might be justified by the fact as Coxhead's list was compiled in 2000, there might have been a big change in the use of adapt from 2000 up to now in which it might have turned to be used mostly in the academic field. Thus, it is suggested that Coxhead's academic list might need to be updated depending on advanced corpora such as COCA.

Given that academic genre and spoken and TV/M genres represent the two extremes of the formality continuum with academic genre being the most formal and spoken and TV/M genre being the least formal. A surprising fact that emerges from the data in table 1 is that the formality continuum is reflected in the distribution of the target verbs in this study in which the highest frequency of roughly all the target verbs concentrates in the most formal genre, namely Academic and the lowest frequency of the target verbs concentrates in the most informal genres, namely Spoken and TV/M. In addition to this, in-between genres such as Magazine and News which may contain both formal and informal forms encompass the middle frequency of the target verbs. To clarify this point, the frequency data of the target verbs in table 1 has been tabulated from highest to lowest. Thus, it becomes immediately apparent that the 1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup> ranks of the highest frequency are distinguished with most formal genres. To exemplify, in the 1<sup>st</sup> rank the highest frequency of all the target verbs is in the Academic genre except for the verb change. The 2<sup>nd</sup> and 3<sup>rd</sup> ranks of highest frequency contains Magazine and Web genres which also represent formal language. Moreover, the approximately absolute absence of informal genres such as Spoken and TV and Movie in the first six ranks confirms the formality continuum in the distribution of the target verbs among all genres. An investigative glance on the last two ranks which display the lowest frequency of the target verbs demonstrates that the lowest frequency occurs in the most informal genres such as Spoken and TV and Movie. Aside from two genres, namely Fiction and Academic, the last two ranks of lowest frequency contain only Spoken and TV and Movie genres. Additionally, the middle ranks. namely 4<sup>th</sup>, 5<sup>th</sup> and 6<sup>th</sup> mostly contain genres such as News, web, and Blog which may contain both formal and informal language.

### **Conclusion**

With respect to the overall frequency of the target verbs, it has been revealed that the target verbs in this study differed from one another in terms of their overall frequency of occurrences in COCA to a great extent. Based on their frequency from highest to lowest, the target verbs are ranked in COCA in this order, *change*, *adjust*, *adapt*, *alter*, *and modify*, respectively.

Of the genre distribution, all the target verbs have the highest presence in written English occurring most frequently in Academic, Magazines, Web, and Blog genres. Additionally, in the other end, the considerably low frequency of the target verbs is simply noticed in *Spoken* genres confirming that the target verbs in this study are marked with high degree of formality. Unlike *adapt, alter, adjust and modify, change* behaved differently in that it scored the highest rank in *Web*, followed by *Blog, Spoken, Magazine, TV* and *Movie, News, Academic* and Fiction. Given that *change* occurred with high frequency in both Spoken and TV and Movie genres which come in the third and fifth ranks, respectively, *change* is

claimed to be informal. This might be enhanced by the fact that academic genre comes in the seventh rank. However, a quick look at the distribution of the frequency of change among the eight genres shows that unlike the other target verbs in this study whose high frequency noticeably concentrates in written English and considerably low frequency concentrates in spoken English, change occurred with markedly high frequency in approximately all genres including both that encompass formal and informal texts. In addition to this, the finding that approximately all the target verbs are associated with academic language is consistent with Coxhead's (2000) Academic Word List (AWL) in which modify, alter, and adjust are listed among the academic words. The absence of change in Coxhead's academic word list is also in line with the findings in this study. Surprisingly, adapt which is absent from Coxhead's academic word list is reported in this study to be most frequent in the academic texts.

What has been mentioned should be enough to evidently prove that the adoption of genre distribution as a criterion to distinguish synonymous words has been demonstrated to be practically effective. Thus, the five synonymous verbs, namely change, alter, adapt, adjust, and modify are not distributed evenly in all genres, signaling that they are near-synonymous verbs that cannot be used interchangeably in all contexts. Thus, it could be implicated that teaching material designers and teachers of English should be convinced to utilize corpora in the classrooms. Teachers could, for example, help their students give preferences to synonyms based on their highest frequency and genre distribution as this significant information is overlooked even by contemporary dictionaries. Further studies may also uncover how the target verbs in this study behave in terms of formality in the British National Corpus (BNC), which is the representative of the British English so the findings in both COCA and BNC might help teaching material designers and teachers of

English acquire an in-depth understanding of the similarities and differences of these near synonyms. The constant availability of and free accessibility of contemporary corpora such as COCA necessitate that teachers should engage the learners in home corpus-based assignments that insure gradual qualification of the learners to explore the language via corpora fruitfully. Additionally, further studies may also use the global English corpus which contains English dialects of different counties such as Indian, Malaysian, and Australian English dialects thus making it possible to compare and contrast the use of the English target verbs in this study worldwide. Of historical changes of the target verbs in this study, future researchers are recommended to employ historical corpora such as the Corpus of Historical American English (COHA). Moreover, the findings of this study can be insightful to researches to target a wider set of frequently used near-synonymous verbs in English other than the target verbs in this study.

#### References

- [1] O'Grady, W., & Archibald, J. (2016). Contemporary Linguistic Analysis: An Introduction Toronto. Pearson.
- [2] Ajmal, M., Kumar, T., Ritonga, M., & Nukapangu, V. (2022). A corpus-based analysis of the adjectives and synonyms: beautiful, handsome, and pretty. *World Journal of English Language*, *12*(2), 159-168.
- [3] AlAmro, M. (2019). A corpus-based study on English synonyms: babble, blather, chatter, gibber, jabber and prattle. *International Journal of Social Sciences & Educational Studies*, 6(1), 122-133.
- [4] Aroonmanakun, V. (2015). Quick or fast: A corpus based study of English synonyms. Language Education and Acquisition Research Network (LEARN) Journal, 8(1), 53-62.
- [5] Baker , P., Hardie, A., & Mcenery, T. (2006). *A glossoary of Corpus Liguistics*. Edinburgh: Edinburgh University Press.
- [6] Boontam, P., & Phoocharoensil, S. (2022). Broaden your horizons: Distribution and

- collocational patterens of the English synonyms expand, widen, and broaden. *The international Journal of Communication and Linguistic Studies*, 20(1), 107-123.
- [7] Cai, J. (2012). Is it "great" enough? A corpusbased study of "great" and its near synonyms. (Master's thesis) Ball State University, Indiana, United State of America.
- [8] Chaengchenkit, R. (2023). A Corpus-based study of the synonyms cease, halt, and stop. *Language Education and Acquisition Research Network*, 16(1), 473-494.
- [9] Chaokongjakra, W. (2023). Is it 'important,' 'significant,' or 'crucial'? A corpus based study of English synonyms. *Language Education and Acquisition Research Network*, 16(2), 512-532.
- [10] Cheng, W. (2012). Exploring Corpus Linguistics: Language in Action. New York: Routledge.
- [11] Chung, Siaw-Fong. (2011). A corpus-based analysis of "create" and "produce". *Chang Gung Journal of Humanities and Social Sciences*, 4(2), 399-425.
- [12] Coxhead, A. (2000). A new academic word list. TESOL Quarterly, 34(2), 213-238. https://doi.org/10.2307/3587951
- [13] Dash, N. S. (2008). *Corpus Linguistics : An Introduction*. New Delhi: Person Education.
- [14] Davies, M. (2020). The corpus of contemporary American English. *English corpora*. https://www.english-corpora.org/coca
- [15] Davies, M. (2010). The Corpus of Contemporary American English as the first reliable monitor corpus of English. Literary and Linguistic Computing: Journal of the Association for Literary and Linguistic Computing, 25(4), 447–464.
- [16] Hu, C., & Yang, B. (2015). Using Sketch Engine to investigate synonymous verbs. *International Journal of English Linguistics*, 5(4), 29-41.
- [17] Hunston, S. (2022). Corpora In Applied Linguistics. New York: Cambridge University Press.
- [18] İŞLER, C. (2022). Near-synonymy analysis of three adverbials: surprisingly, astonishingly and

- amazingly. Urdu Universitesi Sosyal Bilimler Enstitusu Sosyal Bilimler Arastirmalari, 12(2), 627-648.
- [19] Jarunwaraphan, B., & Mallikamas, P. (2020). A corpus-based study of English synonyms: chance and opportunity. *Reflections*, 27(2), 218-245.
- [20] Jones, C., & Waller, D. (2015). *Corpus Linguistics for Grammar : A Guide for Research.* London and New York: Routledge.
- [21] Kayaoglu, N. (2013). The use of corpus for close synonyms. *The Journal of Language and Linguistic Studies*, 9(1), 128-144.
- [22] Kiatthanakul, T. (2015). A corpus based study of English synonyms: achieve, accomplish, and attain. *MA thesis Thammasat University*.
- [23] Kruawonga, T., & Phoocharoensilb, s. (2022). A genre and collocational analysis of the near-synonyms: teach, educate and instruct: A corpusbased approach. *TEFLIN Journal*, *33*, 75-97.
- [24] Lee, C., & Liu, J. (2009). Effects of collocation information on learning lexical semantics for near synonymous distinction. *Computational Linguistics and chinese language processing*, 14(2), 205-220.
- [25] Lertcharoenwanich, P. (2023). Analysis of collocations and semantic preference of the near-synonyms: blank, empty, and vacant. *Language Education and Acquisition Research Network*, 16(1), 365-383.
- [26] Li, E. (2019). A Corpus-assisted study of synonyms in EFL teaching: Take preserve and conserve as examples. *Linguistics and Literature Studies*, 7(2), 39-50.
- [27] Liu, D., & Espino, M. (2012). Actually, genuinely, really, and truly: A corpus-based behavioral profle study of near-synonymous adverbs. *International Journal of Corpus Linguistics*, 17(2), 198-228.
- [28] *Longman Dictionary of Contemporary English* (6<sup>th</sup> ed.). (2014). Person Education Limited.
- [29] Ly, T. H., & Jung, C. K. (2015). A corpus investigation: The similarities and differences of cute, pretty and beautiful. *The Southeast Asian Journal of English Language Studies*, 21(3), 125-140.

- [30] Martinez, R., & Schmitt, N. (2015). Vocabulary. In Biber & Reppen (Eds), The Cambridge Handbook of English Corpus Linguistics (Ist ed., pp.439-459). Cambridge: Cambridge University Press.
- [31] Mikhailov, M., & Cooper, R. (2016). Corpus Linguistics for Translation and Contrastive Studies: A guide for research. New York: Routledge.
- [32] Niwesworakarn, N., Phoocharoensil, S., Petkaew, J., & Intasingh, S. (2023). Collocational patterns of the synonyms: join, participate, and attend: A corpus-based study. *The International Journal of Communication and Linguistic Studies*, 22(1), 63-77.
- [33] O'Keeffe, A., & McCarthy, M. (2010). *The Routledge Handbook of Corpus Linguistics*. New York: Routledge.
- [34] Oxford Advanced Learner's Dictionary (10th ed.). (2015). Oxford Oxford University Press.
- [35] Phoocharoensil, S. (2022). Primary, main, and major: Learning the synonyms through corpus data. *Journal of Language Studies*, 22(4), 76-89.
- [36] Phoocharoensil, S. (2020). A Genre and collocational analysis of consequence, result, and outcome. *The Southeast Asian Journal of English Language Studies*, 26(3), 1-16.
- [37] Phoocharoensil, S. (2021). Semantic prosody and collocation: A corpus study of the near-synonyms persist and persevere. *Eurasian Journal of Applied Linguistics*, 7(1), 240-258.
- [38] Phoocharoensil, S., & Kanokpermpoon, M. (2021). Distinguishing the near-synonyms increase and rise: Genre and collocation investigation. *Kasetsart Journal of Social Sciences*, 42, 968-975.
- [39] Pijuntug, P. (2015). A Corpus-based study on the English synonyms: respond, reply, and answer. *Master's thesis*.
- [40] Sayyed, S. W., & Al-Khanji, R. R. (2019). A corpus-based analysis of eight English synonymous adjectives of fear. *International Journal of Linguistics*, 11(1), 111-138.
- [41]Sinclair, J. (1991). Corpus, concordance, collocation: Describing English language. Oxford University Press.

- [42] Song, Q. (2021). Effectiveness of corpus in distinguishing two near-synonymous verbs: damage and destroy. *English Language Teaching*, 14(7), 8-20.
- [43] Sormet, R. (2017). A corpus-based study of the English synonyms: almost, virtually, and practically. *MA Thesis*, *Thammasat University*.
- [44] Sridhanyarat, K., & Phoocharoensil, S. (2023). A corpus-based investigation of English near synonyms: assess, evaluate and measure. *Humanities, Arts and Social Sciences Studies*, 23(1), 208-219.
- [45] Sriwangrach, B. (2024). A Contrastive corpusbased study of American and English adjectives: important and significant. *International Journal of Education*, 12(2), 32-44.
- [46] Sumonsriworakun, P. (2022). A corpus-based investigation of English synonyms: disadvantage, downside, and drawback. *Language Education* and Acquisition Research Network, 15(2), 649-678.
- [47] Szudarski, P. (2018). Corpus Linguistics for Vocabulary: A guide for research. New York, NY: Routledge.
- [48] Thongpan, N. (2022). A corpus-based study of English synonyms of the adjectives far, distant, and remote. *Journal of Positive School Psychology*, 6(6), 3986-4001.
- [49] Tognini-Bonelli, E.(2001). *Corpus Linguistics at Work. Amsterdam*: John Benjaxnins
- [50] Yuliawati, S., & Indira, D. (2019). Near synonyms and the feature of word sketch difference of the Sketch Engine: A case study of the English adjectives brave and courageous. *International Journal of Recent Advances in Multidisciplinary Research*, 6(3), 4684-4689.
- [51] Yusu, X. (2014). On the application of Corpus of Contemporary American English in vocabulary instruction. *International Education Studies*, 7(8), 68-73.

### Appendix A: the charts of The total frequency and genre distribution of the target verbs in COCA





