

THE CHOICE OF CORPORA FOR CONTRASTIVE ANALYSIS: A CASE STUDY OF THE ENGLISH CAUSATIVE CONSTRUCTION WITH “ENCOURAGE” AND ITS UKRAINIAN COUNTERPART

Iryna Karamysheva

*Lviv Polytechnic National University,
12, S. Bandery Str., Lviv, Ukraine, 79013
iryyna.d.karamysheva@lpnu.ua*

Modern corpus-based contrastive studies brought attention to the choice of *tertium comparationis* in contrastive research. This problem can be overcome by using a bidirectional parallel corpus, although there are concerns about “translation effects” and the lack of proper *tertium comparationis* in parallel corpora.

Corpus-based contrastive analysis of the English-Ukrainian language pair faces the challenge of the lack of representative parallel corpora since they are still in the process of development. The presented research showcases that two separate corpora (COCA and GRAC) can be used to harvest data, if there is a correct choice of *tertium comparationis*, reflected in the query built. The research object is the causative construction with the non-finite complement, triggered with the complement taking predicate “*encourage*”, and its Ukrainian equivalent construction with “*заохомити*”. The parallel corpus ParaRook||EN→UK, nested in GRAC, is used as a complementary in this study. The theoretical background relies additionally on the insights from the Usage-Based Construction Grammar, following the assumption that grammar is shaped by the frequency of use.

Key words: corpus-based contrastive analysis, *tertium comparationis*, causative construction, non-finite complement, Usage-Based Construction Grammar, usage frequency.

Introduction. Over recent years there has been an extensive discussion of the place of Corpus Linguistics in linguistic theory since it has truly revolutionized the world of language study. Recent publications (among others The Routledge Handbook of Corpus Linguistics, 2 ed. (2022) [26], Corpora in Applied Linguistics, 2ed. (2022) [17], The Handbook of Usage-Based Linguistics (2023) [25]), considering the changes and challenges of corpus linguistics over the last decade, focus not only on the methodological progress but also on the academic and social impact of corpus linguistics, highlighting an extensive use of corpora data for researches within a range of linguistic frameworks. The enhanced access to corpora via online interfaces, as remarked by O’Keeffe and McCarthy [23, p. 1], has generally enabled a far broader population to explore data from a greater range of languages than was the case just ten years ago, and the technological progress of corpora themselves have allowed linguists to interpret results that have become more sophisticated.

This paper is meant as a contribution to the corpus-based contrastive studies which is a successful merge of Corpus Linguistics and Contrastive Linguistics. The seminal paper by H. Hasselgård [14, p. 185] describes the development of this new approach towards the contrastive study of languages which was aptly defined by the predecessor of corpus-based contrastive linguistics S. Johansson as “contrastive linguistics in a new key” [18; also 20, p. 437].

Research object and the previous research in the area. Causative constructions are widely studied in Modern English in different frameworks, e.g.: in linguistic typology, pursuing the mainstream generative grammar approach (B. Comrie, 1976 [5]); within the typological study of complementation patterns (M. Noonan, 2007 [22]); within a usage-based approach towards the study of non-finite complementation (Th. Egan, 2008 [11]); combining corpus linguistics and cognitive linguistics (G. Gilquin, 2010 [12]); applying modern collostructional analysis within the Construction Grammar approach (G. Gilquin, 2024 [13]).

The object of this case study is the non-finite complement construction with the meaning of causation, triggered by the verb “*encourage*” in English, for example:

I encourage you to read this handbook thoroughly

(<http://www.smcm.edu/tothepoint/>, 2012) [28].

The following example contains the non-finite complement with the infinitive ‘*you to read*’ and, therefore, is the sentence with secondary predication, containing two types of predication. The primary predication is embodied by a primary subject (S1) and a primary predicate (P1) with causative semantics ‘*encourage*’, triggering the non-finite complement construction which is a secondary predication construction, consisting, in its turn, of secondary subject (S2) typically expressed by a pronoun in the objective case or a noun in the common case and a secondary predicate (P2), expressed by a non-finite (infinitive). Consider this example with the syntactic roles described:

I (S1) encourage (P1) [you (S2) to read (P2) this book (Obj.) thoroughly (adjunct)] [non-finite complement construction].

The mentioned example contains the non-finite complement construction that is simultaneously a secondary predication construction. We deem it proper to make the theoretical provision concerning the use of the term ‘construction’. As it is aptly remarked by H. Boas and M. Fried, representatives of Construction Grammar approach: “[...] the term ‘construction’ is also a very traditional one, used loosely by linguists and non-linguists alike as a descriptive label [...]. It is essential to keep this sense of the term ‘construction’ distinct from the way it is used in Construction Grammar [...]” [4, p. 2]. We follow the understanding of ‘construction’ as suggested by W. Croft who defines ‘construction’ as: “any pairing of form and function in a language (or any language) used to express a particular combination of semantic content and information packaging” [7, p. 17; 6, p. 3]. The packaging of the semantic content can be organized as predication. Consequently, in terms of W. Croft, a sentence containing a non-finite complement construction, will reflect the encoding of one event as the argument of a second event. Only certain predicates allow events as arguments; these predicates are called ‘complement-taking predicates’ or CTPs by M. Noonan [22, p. 53] with this term further used by W. Croft as well [7, p. 551–558]. Following the presented understanding, the primary predicate “*encourage*” is a complement-taking predicate triggering the non-finite complement which is a secondary predication construction itself.

Causative constructions, as exemplified, are also secondary predication constructions. The practical grammars for the Ukrainian learners of English prompt to translate English sentences with such constructions into Ukrainian with the help of subordinate clauses. Little attention is brought to the fact that such constructions do exist in Ukrainian and the reason for this is they are not as productive as English ones. Modern corpora help to showcase the range and productivity of similar Ukrainian constructions. Consider the example of the Ukrainian sentence which contains a similar non-finite complement with the infinitive, triggered by the verb “*заохочити*”:

Моя мама також заохочувала мене читати та займатися мовами (<https://wz.lviv.ua/>, 2012) [30]

Therefore, the main focus of this paper is to carry out the corpus-based contrastive analysis of the English causative construction with “*encourage*” and its Ukrainian counterpart, commenting on the types of corpora suitable for the corpus-based contrastive analysis of the English-Ukrainian language pair.

Methods and data extraction procedure. The main methods applied are *contrastive analysis* and *frequency analysis* since the research is a contribution to corpus-based contrastive grammar studies.

The contrastive analysis proper should start with choosing a correct *tertium comparationis*. Ukrainian linguist N. Andreichuk believes that comparability criterion is one of the key concepts and has to be established prior to any analysis: “The analyst is supposed to answer questions what lingual objects can be compared in the observed languages and what the aspects of comparison are” [2, p. 196]. The representatives of the Norwegian school of corpus-based contrastive studies S. O. Ebeling and J. Ebeling state in a similar key: “It is generally agreed that in order to establish equivalence across languages, a sound *tertium comparationis* is needed, i.e. an objective background of sameness that ensures that we compare like with like. Several *tertia comparationis* have been launched over the years, including surface form, deep structure and translation, but no consensus has been reached...” [8, p.97; more on that 9, p.13–34]. The researchers further argue that: “Within corpus-based contrastive analysis, the types of available *tertia comparationis* are very much tied to the different types of corpora that are typically used in cross-linguistic research, i.e. comparable corpora and parallel corpora of different kinds (e.g., uni-, bi- or multidirectional).” [8, p. 97–98]. In this respect it is worth considering the figure by H. Hasselgård [14, p. 187] that presents types of multilingual corpora for contrastive analysis. The mentioned researchers employ the English-Norwegian Parallel Corpus (ENPC) in their research. It has to be stated that the use of parallel corpora for contrastive analysis causes discussion among the researchers in the field. M. Barlow [3], K. Aijmer and B. Altenberg [1] raise concerns about the availability of “translation effects” in comparable, especially, parallel corpora and the lack of an obvious *tertium comparationis*. H. Hasselgård, a keynote speaker at the latest 10th International Contrastive Linguistics Conference that took place in July 2023, claims that the danger of translation bias in corpus-based contrastive studies can be counteracted by the use of the the bidirectional corpus model. Mutually S. O. Ebeling and H. Hasselgård express an opinion that a bidirectional parallel corpus, which relies on both comparable monolingual and bidirectional translation data, may yield more robust insights into cross-linguistic matters than either of the two on their own [10, p. 30].

Nevertheless, a researcher, currently working with the English-Ukrainian language pair in terms of contrastive analysis, especially a corpus-based one, faces the absence of a large representative parallel corpus, preferably a comparable bidirectional one. M. Shvedova and A. Lukashevskyi, the authors of the recent paper “Creating Parallel Corpora for Ukrainian: a German-Ukrainian Parallel Corpus (ParaRook||DE-UK)”, state that a significant part of the existing Ukrainian parallel corpora is not currently available to the Ukrainian community for various reasons and enumerate which parallel corpora, including the Ukrainian language, are done for particular projects and are available at the moment [24; p.14]. M. Shvedova, one of the creators of GRAC, believes that this Ukrainian language reference corpus is a valuable resource for creating Ukrainian parallel corpora: (M. Shvedova, 2017–2025) since it contains translations from 89 languages, mostly fiction, with a total size of 172 million tokens of texts translated from different languages (GRAC v.17: uacorpus.org). In the mentioned paper M. Shvedova presents ParaRook||DE-UK (Shvedova and Lukashevskyi, 2023–2024), which is the first large German-Ukrainian corpus collected and verified manually, with detailed metaannotation and morphosyntactic annotation, and searchable online. The title refers to the Ukrainian monolingual reference corpus GRAC (grak is the Ukrainian name for rook) and also sounds like “pair of hands” in Ukrainian [24, p.16]. The English-Ukrainian Parallel Corpus ParaRook||EN→UK(2024–2025) is still in the process of its development but is already available on GRAC (<https://www.uacorpus.org/en/poshuk-u-graku/paralelni-korpusi-pararook>).

The presented case study uses two corpora, GRAC [30] and ParaRook||EN→UK [29], for the search of the Ukrainian causative construction, triggered by the CTP “захопити”. The English causative construction with “encourage” is searched with the help of the Corpus of Contemporary American English (COCA), created by Mark Davies, which is by far the most widely-used of the corpora from English-Corpora.org (new version released March 2020) [28].

The case study uses the notion of a ‘construction’, as defined by W. Croft, as *tertium comparationis*. Both W. Croft and H. Boas [7; 4] postulate the notion of ‘construction’ as a comparative concept.

The second method used in this research is *frequency analysis*. In this study we share the claim that “[...] corpus data becomes an indispensable tool for researchers to determine syntactic patterns [...]” [16, p. 18]. The usage-based approach to language proceeds with the assumption that grammar is shaped by the frequency of use and the corpus data play an important role in such studies [15]. The case study takes into account the token frequency produced from the corpus search.

The data extraction procedure from COCA. The causative construction, a studied lingual object, being a non-finite complement construction triggered by the CTP “encourage”, has the following pattern: N1 V N2 non-f V (N3), reflecting the sequence of parts of speech used to express the primary predication in the form of the main clause and the secondary predication as an embedded non-finite complement construction. The pattern predetermines the choice of correct POS tags. To build a proper search query it is, first of all, necessary to take into account the verb (a primary predicate) that serves as a CTP and introduces the non-finite complement construction into the sentence; and, second, the expression of N2 (secondary subject or the semantic subject of the complement construction itself) which is either a pronoun in the objective case or the noun/noun group in the common case. Consequently, two queries have been used:

1) ENCOURAGE PRON TO v?i; 2) ENCOURAGE NOUN TO v?i;

Consider Figure 1 which exemplifies the first query with the examples of sentences obtained (the CTP and the triggered non-finite complement construction is highlighted in bold).

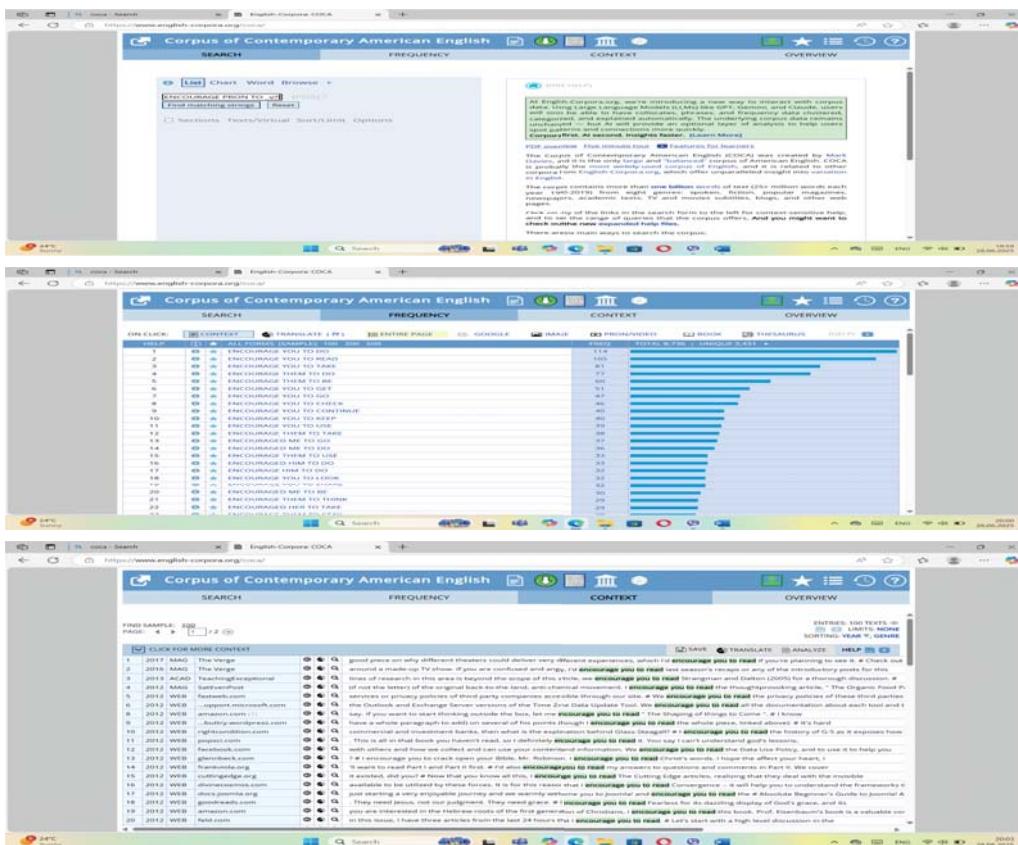


Fig. 1. A combination of screenshots, exemplifying a query in COCA with the examples of sentences obtained.

The data extraction procedure from GRAC. The search procedure with GRAC differs from COCA. The researcher has to use CQL (the Corpus Query Language which is a code used to set criteria for complex searches) for producing a correct query. To search for the equivalent causative construction, triggered in Ukrainian by the CTP “*заохомому*”, the following queries were used, as well taking into account the expression of the secondary predicate either by a pronoun in the accusative case or the noun/noun group in the nominative case:

- 1) [lemma="заохочувати"] [tag=".*pron.*"] [tag=".*inf.*"];
2) [lemma="заохочувати"] [tag=".*noun.*" & tag!=.*pron.*"] [tag=".*inf.*"].

Consider Figure 2 which exemplifies the first query with the examples of sentences obtained (the CTP and the triggered non-finite complement construction is highlighted in red color).

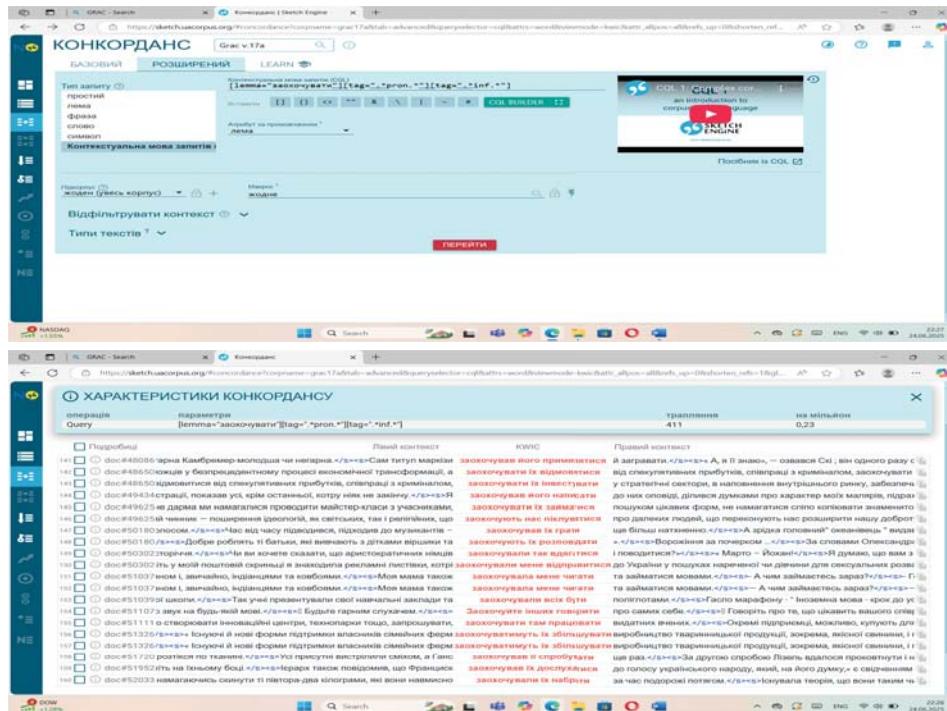


Fig. 2. A combination of screenshots, exemplifying a query in GRAC with the examples of sentences obtained.

The data extraction procedure from ParaRook||EN→UK. The procedure with the parallel corpus presupposes, first of all, the choice of the search direction from English to Ukrainian or from Ukrainian to English, consider Figure 3.

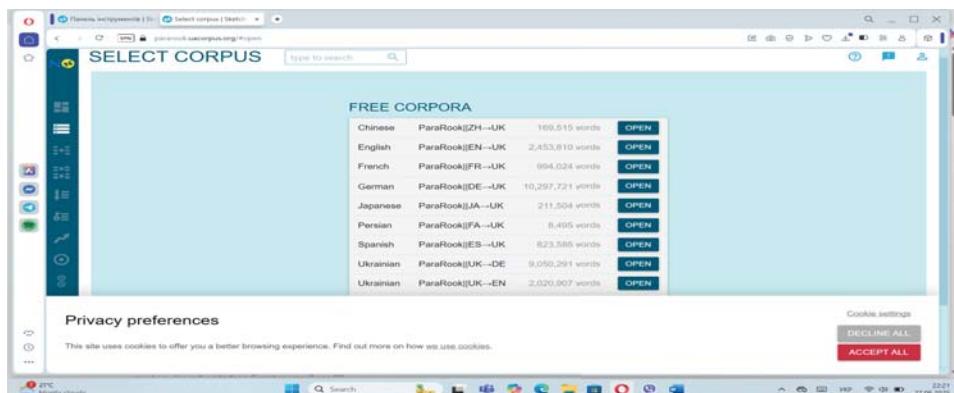


Fig. 3. The interface with the choice of parallel corpora ParaRook within GRAC.

One can search within the English-Ukrainian direction ParaRook||EN→UK, which contains 2,453,810 words at the moment, with a simpler query if one wants a large range of examples: [lemma="encourage"]|[upos="PRON"], see Figure 4:

The top screenshot shows the 'PARALLEL CONCORDANCE' interface for ParaRook||EN→UK. The search query is set to '[lemma="encourage"]|[upos="PRON"]'. The results are displayed in two columns, each showing a snippet of text with the word 'encourage' highlighted in red. The bottom screenshot shows the full list of 30 examples, with the first few entries visible.

Top Screenshot (ParaRook||EN→UK):

```
[lemma="encourage"]|[upos="PRON"]
```

Bottom Screenshot (ParaRook||UK→EN):

Results for the query [lemma="encourage"]|[upos="PRON"] (30 examples):

- <> They were both willing to come with me, but I **encouraged them** to go have fun — they didn't know how preoccupied I could get when surrounded by books; it was something I preferred to do alone. </>
- <> Whatever he saw there must have **encouraged him** . </>
- <> This cheerful way of looking at the matter **encouraged me** to dare sudden expression of a long covetousness. </>
- <> In my first year, I was not skilled enough to make any of the teams. </><> But during my second year, my friend Locke Nozamella, Headtown's champion hurdler, **encouraged me** to take up a new sport: long-distance running. </>
- <> Like the regent, he saw my future role as counselor to Sabata, and he **encouraged me** to study law. </>
- <> They wanted the chiefs to **encourage their** subjects to come to the Reef. </>
- <> From the magistrate's office, we went straight to the township where the demonstration was taking place, and even from half a mile away we heard the robust singing of our volunteers and the great crowd of supporters who had come to **encourage them** . </>
- <> Mr. Basner was an excellent boss and as long as I got my work done at the firm he **encouraged my** political work. </>
- <> I was keen that he should go and **encouraged him** to do so, whether or not he conferred with the executive. </>

Fig. 4. A combination of screenshots, exemplifying a query in ParaRook||EN→UK with the examples of sentences obtained.

Otherwise the researcher can choose the Ukrainian-English direction ParaRook||UK→EN, which contains 2,020,907 words, and search with a more specialized query to get the exact construction studied.

Figure 5 exemplifies the search with a query [lemma="заохотити"] [upos="PRON" & morphology="Case=Acc.*PronType=Prs.*"] [morphology=".*VerbForm=Inf.*"] and the example obtained.

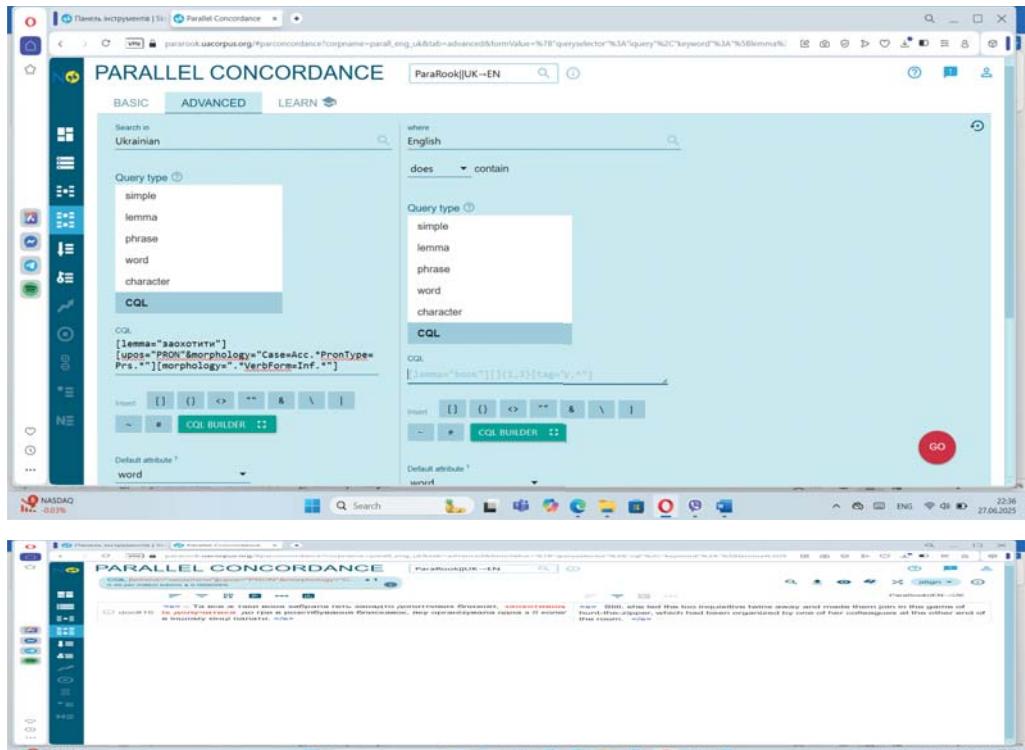


Fig. 5. A combination of screenshots, exemplifying a query in ParaRook||UK→EN with the example obtained.

Results and Discussion. In this research, we follow the claim of the usage-based approach to the language study that frequency of usage is important for the description of construction functioning. Modern corpora give the possibility to obtain the frequency of the chosen lingual object under study.

Data for the case study of the English causative construction, triggered by the CTP “encourage”, were extracted from the COCA. This construction appeared to be the fifth in a frequency row among the English causative constructions with the Infinitive within the non-finite complement part. Table 1 presents these data.

Table 1

Frequency data of the most frequent English causative constructions,
containing non-finite complement with the Infinitive

	CTP, triggering the construction	Total frequency
1	MAKE PRON _v?i	84,182 tokens
	MAKE NOUN _v?i	4,359 tokens
2	ASK PRON TO _v?i	40,834 tokens
	ASK NOUN TO _v?i	7,689 tokens
3	FORCE PRON TO _v?i	16,357 tokens
	FORCE NOUN TO _v?i	5,025 tokens
4	CAUSE PRON TO _v?i	10,418 tokens
	CAUSE NOUN TO _v?i	4,234 tokens
5	ENCOURAGE PRON TO _v?i	8,736 tokens
	ENCOURAGE NOUN TO _v?i	7,200 tokens

Causative constructions are the subtype of the manipulative constructions, and as it is stated by M. Noonan: “Manipulative CTPs typically encode situations where the agent attempts to manipulate the affectee into performing some action or assuming some state” [22; p. 136]. He further argues that manipulative predicates may be simple (“cause”) or may in addition encode information about the manner of causation (“force”, “make»], etc.). Taking into account the semantics of the CTP “encourage”, given by the Cambridge Dictionary online [27] (to make someone more likely to do something, or to make something more likely to happen; 2) to help someone to feel confident and able to do something, or to give advice to someone to do something), this primary predicate is expressing mild manipulation, largely in the form of support and advice. Comparing the frequency data of the expression of the secondary subject by the pronoun in the objective case or by the noun in the common case, it becomes obvious that for “encourage” these frequencies do not differ drastically while with other CTPs, especially with “make”, the use of the pronoun in the objective case considerably prevails.

Data extraction of the Ukrainian equivalent causative construction, triggered with the CTP “заохочувати” from the corpus GRAC. This construction also appeared to be the fifth in a frequency row among the Ukrainian causative constructions with the Infinitive within the non-finite complement part, although the list is not exactly the same. Table 2 presents these data.

Table 2

**Frequency data of the most frequent Ukrainian causative constructions,
containing non-finite complement with the Infinitive**

	CTP, triggering the construction	Total frequency
1	[lemma="попросити просити"] [tag=".*pron.*"] [tag=".*inf.*"]	35,683 tokens
	[lemma="попросити просити"] [tag=".*noun.*" &tag!=“.*pron.*”] [tag=".*inf.*"]	43,449 tokens
2	[lemma="примусити змусити"] [tag=".*pron.*"] [tag=".*inf.*"]	24,100 tokens
	[lemma="примусити змусити"] [tag=".*noun.*" &tag!=“.*pron.*”] [tag=".*inf.*"]	21,661 tokens
3	[lemma="порекомендувати рекомендувати"] [tag=".*pron.*"] [tag=".*inf.*"]	1,823 tokens
	[lemma="порекомендувати рекомендувати"] [tag=".*noun.*" &tag!=“.*pron.*”] [tag=".*inf.*"]	6,927 tokens
4	[lemma="спонукати"] [tag=".*pron.*"] [tag=".*inf.*"]	3,097 tokens
	[lemma="спонукати"] [tag=".*noun.*" &tag!=“.*pron.*”] [tag=".*inf.*"]	3,659 tokens
5	[lemma="заохочувати"] [tag=".*pron.*"] [tag=".*inf.*"]	0,418 tokens
	[lemma="заохочувати"] [tag=".*noun.*" &tag!=“.*pron.*”] [tag=".*inf.*"]	1,067 tokens

Comparing the frequency data of the expression of secondary subject by the pronoun in the accusative case or by the noun in the nominative case, it becomes obvious that, unlike in English, we have more nouns as secondary subjects after the Ukrainian CTPs. The same is true for the CTP “заохочувати”.

Taking data for contrastive analysis from different corpora presupposes the normalization of the frequency data per million in order to compare the obtained figures. The corpus COCA contains more than one billion words whereas the Ukrainian GRAC (version 17) has 1,781 milliard tokens. Therefore, the normalized data for the studied English construction with “encourage” and its Ukrainian counterpart with “заохочувати” will be the following (the expression of the secondary subject by the pronoun and the noun is taken together), consider Table 3.

Table 3

**The quantity of constructions from COCA and GRAC
given as their relative frequency per million**

Total frequency per COCA	Relative frequency per million	Total frequency per GRAC	Relative frequency per million
15,936 tokens	15,936	1,485 tokens	0,83

Consequently, the higher productivity of the English constructions is obvious in comparison to the Ukrainian ones. At the same time, the data from the Ukrainian corpus prove that such constructions do exist in Ukrainian and the rendering of the English secondary predication constructions of such a type should not be done entirely with subordinate sentences.

The experiment search with the parallel corpus ParaRook did not produce numerous results because these corpora are still in the process of enlargement; nevertheless, the examples

are vivid to produce the preliminary conclusions. The search with the ParaRook||UK→EN produced only one example, matched with the following original sentence:

Та все ж таки вона забрала геть занадто допитливих близнят, заохочивши їх долучитися до гри в розстібування блискавок, яку організувала одна з її колег в іншому кінці палати. (FIC: Aldous Huxley. Brave New World. / Олдос Гакслі. Який чудесний світ новий! Пер. Віктор Морозов, 2016) [29].

Still, she led the too inquisitive twins away and made them join in the game of hunt-the-zipper, which had been organized by one of her colleagues at the other end of the room [29].

The English causative construction triggered by the CTP “made” is rendered with the milder form of manipulation by the synonymous “encourage” – “заохочити” in Ukrainian.

The search with the ParaRook||EN→UK produced 30 examples. Out of 30 examples only 12 are causative constructions with the non-finite complement, being simultaneously secondary predication constructions. Table 4 exemplifies the ways of rendering English constructions into Ukrainian.

Table 4

Examples of rendering the English causative construction with the CTP “encourage” into Ukrainian

English causative constructions, triggered with “encourage”	Ukrainian equivalents	The way of rendering of English constructions into Ukrainian
<i>... encouraged them to drink and have another ...</i>	<i>... укоєськав їх допити віск ...</i>	with the equivalent non-finite complement but a synonymous CTP
<i>... a kind-voiced man was patting me and encouraging me to rise.</i>	<i>Якийсь чоловік плескав мене по граві й лагідно просив підвестися.</i>	with the equivalent non-finite complement but a synonymous CTP and omission of the secondary subject
<i>We don't encourage them to indulge in any solitary amusements.</i>	<i>Ми не заохочуємо їх віддаватися самотнім умікам.</i>	with the equivalent non-finite complement construction
<i>I encouraged them to go have fun ...</i>	<i>Я переконала їх іти розважатися ...</i>	with the equivalent non-finite complement but a synonymous CTP
<i>My friend ... encouraged me to take up a new sport.</i>	<i>Мій друг...порадив мені спробувати нову дисципліну.</i>	with the equivalent non-finite complement but a synonymous CTP
<i>This cheerful way of looking at the matter encouraged me to dare ...</i>	<i>Його слова додали мені сміливості ...</i>	descriptive translation
<i>... he encouraged me to study law.</i>	<i>... заохочував вивчати право...</i>	with the equivalent non-finite complement construction but omission of the secondary subject
<i>I ...encouraged him to do so.</i>	<i>Я ... заохочував його так і зробити.</i>	with the equivalent non-finite complement construction
<i>The sight of freedom looming on the horizon should encourage us to redouble our efforts.</i>	<i>Образ свободи, що вже мріє на виднокраї, має надихнути нас на те, щоб подвоїти зусилля.</i>	descriptive translation
<i>... this would encourage him to go further</i>	<i>... і таким чином надихнути його на дальші кроки.</i>	descriptive translation
<i>We've got to encourage them to evolve!</i>	<i>Ми повинні заохочити їх до розвитку!</i>	descriptive translation
<i>... an attempt ... encourage them to take responsibility...</i>	<i>це була спроба ... прихилити їх брати відповідальність ...</i>	with the equivalent non-finite complement but a synonymous CTP

The analysis of these examples shows that we have translations with the equivalent non-finite complement construction, triggered with the Ukrainian CTP “*заохомити*”. Some translations are with the equivalent non-finite complement but a synonymous CTP and the part of examples are descriptive translation.

There was made a similar attempt for the search of a “*make*”-construction and its Ukrainian counterpart in the experimental parallel corpus, developed under the scientific supervision of Nataliia Kotsyba (available at: <https://mova.institute>; the number of tokens of the English part is 1.5 million.) [19]. The analysis of the obtained results showed that only in a few cases an English construction is translated by its direct Ukrainian equivalent. The majority of matched examples were a certain type of paraphrase, which are more suitable to study strategies used in the process of translation. The benefits and application of parallel corpora, primarily for the purposes of Translation Studies, are described in [21].

Consequently, we argue that it is more beneficial to use separate representative corpora for the purposes of the contrastive analysis research within the English-Ukrainian language pair. In this case study two large representative corpora were used – COCA and GRAC that have substantial volume of data, and the lingual object in focus can be researched also across registers. The parallel corpora can be used for complementary purposes unless they are developed according to a bidirectional comparable model, similar to an English-Norwegian Parallel corpus, suitable for the purposes of contrastive analysis.

Conclusions. This case study focused on the corpus-based contrastive analysis of the English causative construction, triggered with the CTP “*encourage*” and its Ukrainian counterpart. The mentioned causative constructions contain the non-finite complement with the Infinitive, being simultaneously secondary predication constructions within the sentence. The novelty of this research is the corpus-based and usage-based approach towards the analysis of such constructions as well as the contrastive study of their Ukrainian equivalents which were not studied in a similar way.

The main aim of the paper was to showcase that two separate representative corpora are suited for the corpus-based contrastive analysis provided there is a proper choice of *tertium comparationis*. The notion of ‘construction’ is taken as the platform for comparison and is understood within the Construction Grammar approach which helped to build the correct queries for the data extraction.

The obtained results show that at the present stage of parallel corpora development it is expedient to carry out the contrastive study of lingual objects within the English-Ukrainian language pair, using two separate representative corpora. Data harvested in such a way will provide the researcher with the clear picture of the peculiarities of functioning of the studied lingual object, supplying the frequency of use and helping to conclude about the productivity of the researched unit. The study of the range of causative constructions with the non-finite complement in English by contrasting them with the equivalent Ukrainian constructions while applying usage-based approach will outline the prospects of further research.

REFERENCES

1. Aijmer K., Altenberg B. (eds). *Advances in Corpus-based Contrastive Linguistics. Studies in Honour of Stig Johansson*. Amsterdam : Benjamins, 2013. P. 301–306.
2. Andreichuk N. *Contrastive Linguistics: Study manual*. Lviv : Ivan Franko National University of Lviv Publishing Center, 2015. 343 p.
3. Barlow M. Parallel texts and corpus-based contrastive analysis. *Current Trends in Contrastive Linguistics: Functional and cognitive perspectives* / Ed. by González, María de los Ángeles Gómez, et al. John Benjamins Publishing Company, 2008. P. 101–121.
4. Boas H. C., Fried M. Introduction. *Grammatical Constructions: Back to the roots (Constructional Approaches to Language 4)* / M. Fried, H. C. Boas (eds.). Amsterdam : John Benjamins, 2005. P. 1–9.
5. Comrie B. The syntax of causative constructions: cross-language similarities and divergences. *Syntax and Semantics*. Vol. 6. *The Grammar of Causative Constructions* / Ed. M. Shibatani). New-York, San Francisco, London : Academic Press, 1976. P. 261–312.
6. Croft W. Constructions across and within languages. *Book of Abstracts ICCG 13. The 13th International Conference on Construction Grammar*. Göteborg, 26–28 August 2024. P 3–4.
7. Croft W. *Morphosyntax: Constructions of the World's Languages*. Cambridge University Press, 2022. 688 p.
8. Ebeling S. O., Ebeling J. Contrastive Analysis, tertium comparationis and corpora. *Nordic Journal of English Studies (NJES)*. 2020. (19). P. 97–117.
9. Ebeling S. O., Ebeling J. *Patterns in Contrast*. Amsterdam : John Benjamins, 2013. P.13–34.
10. Ebeling S. O., Hasselgård H. Combining mutual correspondence and textual variation in a study of English GO and Norwegian GÅ. *Nordic Journal of English Studies (NJES)*. 2024. 23(2). P. 6–33.
11. Egan Th. Non-finite Complementation. A Usage-based Study of Infinitive and -ing Clauses in English. Amsterdam-New York, 2008. 432 p.
12. Gilquin G. *Corpus, Cognition and Causative Constructions. Studies in Corpus Linguistics*, V. 39. Amsterdam, Philadelphia : John Benjamins, 2010. 328 p.
13. Gilquin G. Transfer of collostructions: the case of causative constructions. *Corpus Linguistics and Linguistic Theory*, 2024. P. 1–21.
14. Hasselgård H. Corpus-based contrastive studies: Beginnings, developments and directions. *Languages in Contrast: International Journal for Contrastive Linguistics*. 2020. 20 (2). P.184–185.
15. Horsch J. From corpus data to constructional networks: Analyzing language with the Usage-based Construction Grammar framework. *Journal of Linguistics*. 2023a. Vol. 74, No. 3.P. 701–740.
16. Horsch J. The comparative correlative construction in World Englishes: a usage-based construction grammar approach. Eichstätt : Catholic University of Eichstätt-Ingolstadt. 2023b. PhD thesis.
17. Hunston S. *Corpora in Applied Linguistics*. Second edition. Cambridge, United Kingdom; New York, NY : Cambridge University Press, 2022. 341 p.
18. Johansson S. *Seeing through Multilingual Corpora. On the Use of Corpora in Contrastive Studies*. *Studies in Corpus Linguistics*. 26. Amsterdam : John Benjamins, 2007. P. xiii-xiv.

19. Karamysheva I. Challenges of corpus-based contrastive studies: a case study of the English “make”-construction and its Ukrainian counterpart. *2024 IEEE 19th International Conference on Computer Science and Information Technologies (CSIT)*, Lviv, Ukraine, 2024. P. 1–4.
20. Karamysheva I. *Tertium Comparationis* in contrastive studies: theoretical insights and current trends. *Modern Philology: Theory, History, Methodology*: Scientific monograph. Riga, Latvia : “Baltija Publishing”, 2024. P. 427–446.
21. Matvieieva S., Lemish N., Zernetska A., Babych V., Torgovets M. English-Ukrainian parallel corpus: Prerequisites for building and practical use in Translation Studies. *Studies about Languages, Kalbų studijos*. 2022. № 40. P. 61–74.
22. Noonan M. Complementation. *Language Typology and Syntactic Description*, vol. II: *Complex Constructions*, 2nd edition / ed. T. Shopen. 2007. P. 52–150.
23. O’Keeffe A., McCarthy M. J. ‘Of what is past, or passing, or to come’: corpus linguistics, changes and challenges. *The Routledge Handbook of Corpus Linguistics* 2 ed., / ed. by A. O’Keeffe, M. J. McCarthy. London and New York : Routledge, 2022. P. 19.
24. Shvedova M., Lukashevskyi A. Creating parallel corpora for Ukrainian: A German-Ukrainian parallel corpus (ParaRook||DE-UK). *Proceedings of the Third Ukrainian Natural Language Processing Workshop (UNLP) @ LREC-COLING 2024*. Torino, Italia. ELRA and ICCL. P. 14–22.
25. The Handbook of Usage-Based Linguistics. Ed. by M. Díaz-Campos, S. Balasch. Hoboken, New Jersey : Wiley-Blackwell, 2023. 601 p.
26. The Routledge Handbook of Corpus Linguistics, 2 ed. / Ed. by A. O’Keeffe, M. J. McCarthy. London and New York : Routledge, 2022. 692 p.

SOURCES

27. Cambridge Dictionary. URL: <https://dictionary.cambridge.org/>
28. Corpus of Contemporary American English. URL: <https://www.english-corpora.org/coca/>
29. Parallel Corpora for Ukrainian (ParaRook). URL: <https://www.uacorpus.org/en/poshuk-u-graku/paralelni-korpusi-pararook>
30. Shvedova M., Waldenfels R. von, Yarygin S., Rysin A., Starko V., Nikolajenko T., Lukashevskyi A. et al. (2017-2025): GRAC: General Regionally Annotated Corpus of Ukrainian. Electronic resource : Kyiv, Lviv, Jena. URL: <https://www.uacorpus.org>.

Article submitted 01.07.2025
Accepted for publication 01.09.2025

ВИБІР КОРПУСІВ ДЛЯ КОНТРАСТИВНОГО АНАЛІЗУ: ПРИКЛАД АНГЛІЙСЬКОЇ КАУЗАТИВНОЇ КОНСТРУКЦІЇ, ЩО ВВОДИТЬСЯ ПРЕДИКАТОМ “ENCOURAGE”, ТА ЇЇ УКРАЇНСЬКОГО ВІДПОВІДНИКА

Ірина Карамишева

Національний університет “Львівська політехніка”,
вул. С. Бандери, 12, Львів, Україна, 79013
iryuna.d.karamysheva@lpnu.ua

Корпусна лінгвістика тісно поєдналася з іншими галузями лінгвістики, спричинивши, зокрема, появу корпуснобазованих зіставних досліджень. Це привернуло увагу до вибору та правильного застосування *Tertium Comparationis* у корпуснобазованих контрастивних дослідженнях. Представники норвезької школи корпусних зіставних досліджень аргументують, що проблему з вибором *Tertium Comparationis* можна подолати, використовуючи двонаправлений паралельний корпус, який містить як монолінгвальні дані, так і двонаправлені дані на основі перекладених текстів; інші дослідники в цій галузі досі попереджають про “ефекти перекладу” та відсутність належного *Tertium Comparationis* у паралельних корпусах.

Дослідник, який бажає здійснити контрастивний аналіз англійсько-української мовою пари на основі корпусів, наштовхується на відсутність репрезентативних паралельних корпусів, оскільки вони все ще перебувають у процесі розробки. Головною метою представленого дослідження є демонстрація того, що два окремі корпуси (СОСА та ГРАК) можуть бути використані для збору даних для контрастивного аналізу за умови правильного вибору *Tertium Comparationis*, що відображається у побудованому пошуковому запиті.

У дослідженні поняття “конструкції” використано як платформу для зіставлення, а розуміння “конструкції” прийнято у руслі теоретичної концепції конструкційної граматики. Об’єктом дослідження є каузативна конструкція з нефінітним комплементом, яка вводиться в англійській мові дієслівним предикатом “*encourage*”, та її українська еквівалентна конструкція, що вводиться предикатом “*заохочувати*”. Речення, що містять такі конструкції відносимо до речень з вторинною предикацією. Процес дослідження передбачав побудову правильних пошукових запитів, що формуються по різноманітному в обох корпусах. Сформовані запити дали змогу зробити вибірку таких конструкцій в англійській та українській мовах. Отримані дані співставлено з більш частотними конструкціями такого типу та зроблено висновки щодо їх продуктивності у зіставлювальних мовах. Паралельний корпус ParaRook||EN→UK, що розміщений у ГРАК, використано як додатковий у цьому дослідженні. Приклади речень, що містять досліджувані конструкції в англійській мові, проаналізовано з огляду на отримані відповідники українською мовою, поміж яких засвідчено як еквівалентні конструкції, так і речення з описовим перекладом.

Теоретична основа цього дослідження спирається на засади ужитковобазованої граматики конструкцій, де корпусні дані та частота вживання досліджуваних конструкцій відіграють важливу роль.

Ключові слова: корпуснобазований контрастивний аналіз, *Tertium Comparationis*, каузативна конструкція, нефінітний комплемент, ужитковобазована граматика конструкцій, частота вживання.