Maryana **TOMENCHUK**

USING THE CORPUS APPROACH IN TEACHING ENGLISH IDIOMS

Abstract. Idioms, as cultural reflections deeply ingrained within language, offer invaluable insights into social values, customs, and beliefs. This article studies the crucial role that idioms play in foreign language teaching, focusing on the powerful tool of the Corpus of Contemporary American English (COCA) to enhance language proficiency and cultural understanding.

The study on incorporating corpus linguistics to teach idioms in EFL classrooms holds significant relevance as it enriches language skills, fosters intercultural awareness, and empowers language instructors to make informed decisions for effective teaching.

The aim of this research is to expand the existing literature by introducing an algorithm and instructional activities for teaching idioms through corpus analysis, offering valuable insights to language educators striving to optimize their students' learning experiences with idiomatic expressions.

The methodology includes guiding students in accessing and navigating the corpus, conducting targeted searches for idioms, examining their variations and contextual usage, and engaging in data-driven learning to enhance their comprehension and usage of idiomatic expressions.

Immersing students in contextualized idiomatic expressions across diverse genres, registers, and historical periods, teachers can cultivate a profound appreciation for the dynamic evolution of language over time. The significance of contextual analysis is underscored, guiding educators in enabling learners to discern the subtle nuances and shades of meaning in idiom usage. Engaging students in creative exercises, such as composing dialogues and stories, can foster practical application and reinforce the seamless integration of idioms into authentic language contexts. This article emphasizes the thoughtful usage of the COCA corpus as an essential resource in EFL classrooms by teachers who strive to help their students in acquiring the rich tapestry of idioms and gaining a deeper understanding of language and culture.

Keywords: idiom, corpus, corpus linguistics, COCA corpus, EFL teaching.

Problem statement. The use of corpora in language teaching has ushered in a transformative era in applied linguistics. The field of corpus linguistics harnesses computer technology to analyse extensive sets of spoken and written texts, known as corpora, enabling researchers to gain valuable insights into language usage patterns. With the increasing interest in corpus-based studies, there is growing recognition of the potential of corpus linguistics in revolutionizing language instruction, particularly in the context of second and foreign language education [Conrad 2000].

Prior to the advent of electronically stored corpora, analysing language usage patterns, especially idiomatic expressions, was a laborious task due to the sheer volume of data. However, with the development of computer programs, automated language analysis has become feasible, leading to significant findings in the study of idioms and their role in language communication. Corpus linguistics research has provided language teachers with valuable frequency data, enabling them to determine the significance of idiomatic expressions within language classrooms. Additionally, corpus analysis has shed light on the interdependence of vocabulary and grammar in meaning construction, transforming the conventional understanding of word knowledge.

Despite the promising potential of corpus linguistics in language teaching, there are challenges to address. First, there is a need to develop effective algorithms and instructional activities to effectively integrate corpora into the classroom for idiom teaching. Second, there is a lack of clarity regarding the frequencies and exposure learners gain with idiomatic expressions, which may hinder effective idiom learning. Third, the use of corpus data in language teaching is not without criticism, with concerns about its applicability and potential confusion for students. To address these challenges, this research will focus on exploring the use of the COCA (Corpus of Contemporary American English) corpus for teaching idioms in EFL classrooms.

Analysis of the previous research. The integration of corpus linguistics into language teaching has gained significant attention in recent years, as researchers explore its potential to revolutionize language education. In this literature review, we focus on academic studies related to the use of corpora to enhance the teaching of English as a foreign language.

S. Conrad presents a compelling argument for the transformative impact of corpus linguistics on language instruction [Conrad 2000]. His work highlights the potential of corpus-based insights to fundamentally alter the design of learning materials and curricula. This seminal perspective sets the foundation for subsequent research exploring the practical application of corpus analysis in language classrooms.

D. Biber and R. Reppen contribute to the discussion by emphasizing the significance of corpus-derived frequency data for language teachers [Biber & Reppen 2002]. Their research demonstrates how corpus-based frequency lists can aid in identifying the most commonly used structures in English. Such data-driven approaches offer valuable guidance to educators seeking to prioritize certain expressions within language classrooms.

The interplay between vocabulary and grammar in meaning construction is a crucial aspect of corpus linguistics research, as emphasized by S. Hunston [Hunston 2002]. The work reveals that a word meaning is expressed through its association with other words and syntactic structures. This insight challenges traditional definitions and encourages the adoption of corpus analysis tools to explore collocation and phraseology. It is within this context that idioms gain significance as vital components of linguistic competence.

G. Kennedy provides valuable insights into the different approaches to idiom learning, highlighting the significance of context and repetition [Kennedy 2008]. The idea of avoiding rote memorization gains traction, as learners are encouraged to acquire idiomatic expressions implicitly through repeated exposure and contextual inference.

The role of context in comprehending idiomatic expressions is further explored by S. Webb [Webb 2007] and F. Boers and S. Lindstromberg [Boers & Lindstromberg 2008]. Their research underscores the importance of using diverse contexts to enhance vocabulary learning, allowing learners to grasp the nuanced meanings of idioms. This aligns with the idea of using the corpus to examine idiomatic expressions in various registers and genres, as proposed in this article.

The academic literature reviewed here highlights the growing interest in incorporating corpus linguistics into language teaching to enhance idiom learning in EFL classrooms. These studies focus on the potential benefits of data-driven instruction, contextual analysis, and dynamic digital resources for effective idiomatic expression acquisition. However, they also reveal the need to address certain challenges, such as the lack of clarity on exposure frequencies and the risk of confusion for learners.

The research on using corpus linguistics to teach idioms in EFL classrooms is highly **relevant** as it enhances language proficiency, fosters cross-cultural understanding, and empowers language teachers to make data-driven decisions for effective instruction. Additionally, it encourages active student engagement and contributes to advancements in language education practices.

Research aim and objectives. This article aims to contribute to the existing body of literature by presenting an algorithm and instructional activities for teaching idioms using corpus analysis, providing practical insights for language educators who seek to optimize idiomatic expression learning experiences for their students.

The methodology of the study focuses on presenting the algorithm and instructional activities for teaching idioms. It involves using corpus linguistics as a valuable resource to study real-world language usage. The methodology entails guiding students to access and navigate the corpus, conduct targeted searches for idioms, explore their variations and contextual usage, and engage in data-driven learning to enhance their understanding and usage of idiomatic expressions.

Presentation of the basic material. Corpus linguistics is a field within applied linguistics that employs computer technology to examine extensive sets of spoken and written texts, known as corpora. These corpora are meticulously crafted to represent specific areas of language usage, such as informal conversations or academic writing.

The use of corpora in language teaching

Over the years, researchers have shown a growing inclination towards using the insights derived from corpus-based studies in the realm of second and foreign language education. In fact, some experts, like S. Conrad [Conrad 2000], have put forth the notion that corpus linguistics has the potential to revolutionize language instruction by fundamentally altering our approach to designing learning materials and developing curricula.

Before the advent of electronically stored corpora, it was impractical to identify language usage patterns, such as those found in American English conversations, due to the immense task of manually analysing millions of words within a reasonable timeframe. However, with the development of computer programs, automated language analysis has become possible, leading to corpus-based findings that both enhance and challenge our previous understanding of language use. In particular, the past few decades have witnessed a surge in available information pertaining to (a) the frequency of specific words or linguistic features in a language, (b) the interplay between vocabulary and grammar in meaning construction, and (c) the influence of situational factors, such as communication mode and purpose, on the choices made by language users as writers and speakers.

Corpus linguistics has provided language teachers with a vast amount of frequency data, which can greatly assist them in determining what is more or less significant within a language classroom based on actual language usage patterns [Biber & Reppen 2002, p. 203]. For instance, if teachers want to identify the most commonly used phrasal verbs in English, they can refer to corpus-based frequency lists. When preparing university students for the linguistic demands of academic study in English-speaking institutions, teachers now have access to lists that highlight the most frequently employed academic words across various disciplines, along with significant multiword phrases prevalent in university spoken and written communication.

Through corpus linguistics research, it has been revealed that vocabulary (lexis) and grammar are interdependent in conveying meaning. This understanding has fundamentally transformed our perception of what it truly means to "know" a word. Instead of solely relying on word definitions, dictionaries and language teaching materials are increasingly incorporating corpus analysis tools to offer additional insights into collocation (the tendency of words to co-occur with others) and phraseology (the occurrence of words in specific grammatical patterns) [Hunston 2002]. It is now evident that a word cannot be comprehensively defined in isolation; rather, its meaning is expressed through its association with other words and syntactic structures. Concordancing programs, among other corpus analysis tools, enable teachers and students to examine multiple examples of a given keyword within its contextual framework, facilitating the study of how context influences word meaning. These tools prove particularly valuable for exploring various senses of a word, subtle distinctions between seemingly synonymous words,

and the positive or negative connotations conveyed by specific phraseological expressions [Sinclair 2005].

Corpus linguistics has also highlighted to teachers the significance of register and how language usage varies based on the communicative requirements of a specific situation. Registers can be broadly defined (e.g., comparing writing as a general register to speaking) or more specialized (e.g., studying the register of editorials in the New York Times). As S. Conrad points out, it is no longer effective to provide students with a singular, unified description of the language they are learning. The choices we make in terms of word selection, phrases, and grammatical structures are heavily influenced by the communicative context [Conrad 2000, p. 552]. For instance, the time constraints of conversation can account for the frequent use of contractions and pronouns, while the characteristics of written language (e.g., planning time, the need for contextualization) explain the frequent use of complex noun phrases. Given that language learners encounter various registers in their target language, such as informal conversations with friends, attending university classes, or delivering work presentations, it is crucial for teachers to comprehend how language usage differs across these contexts. It is important for teachers to emphasize significant register variations in their classroom instruction.

Teachers, keen on integrating corpus linguistics into their classrooms, have various options at their disposal. Even if they are new to the field or face limitations in time and resources for corpus activities, they can still benefit from a range of corpus-based resources when preparing lessons and crafting teaching materials. These resources encompass corpus-based dictionaries, grammars, and language textbooks [McCarthy 2008]. Additionally, teachers can adopt a discerning approach when evaluating existing language teaching materials, considering the aspects that corpus linguists have emphasized as significant. These include paying attention to the frequency of word usage when selecting target vocabulary, going beyond word definitions by teaching collocations and syntactic patterns, and highlighting the differences in language use between different registers such as spoken and written contexts.

Several online tools enable language teachers to analyse text collections they intend to use in their classrooms. These tools assist in evaluating the text difficulty and exploring how vocabulary words are utilized throughout an instructional unit [Cobb 2007]. Moreover, teachers can access various online corpora when preparing their lessons and materials. For instance, the Corpus of Contemporary American English (COCA) at www.americancorpus.org permits teachers to not only search for specific keywords but also find synonyms, collocates, grammatical patterns, and idioms. Similarly, the Michigan Corpus of Academic Spoken English (MICASE) empowers users to generate concordance lines and download complete transcripts of speech events such as class lectures, student presentations, and study group sessions.

In addition to using corpora themselves, teachers also have the option of introducing corpora to their students. L. Flowerdew suggests that students can greatly benefit from explicit instruction on how to utilize a wide range of corpus-based resources to investigate their own language-related inquiries [Flowerdew 2009]. Teachers can also involve students in what T. Johns refers to as "data-driven learning", which involves analysing concordance lines that have been selected, organized, and potentially modified by the teacher to draw learners' attention to language usage patterns [Johns 2000]. Expanding on Johns's concept, several researchers have advocated for a "discovery learning" approach, where students, rather than teachers, assume primary responsibility for searching and analysing corpus data. These researchers argue that discovery learning, in contrast to teacher-directed activities, empowers learners to question textbook explanations and native-speaker intuitions and allows them to make language use discoveries that would not have been possible without the use of corpora [Aston 2001; Bernardini 2001].

While the use of corpus data in language teaching has gained popularity, it is not without its critics. Some individuals have expressed scepticism regarding the applicability of corpora in language instruction. There are concerns that divorced from their original context, corpus data may actually confuse students rather than facilitate their learning. To address these concerns, the use of learner corpora [Granger 2002] and small, specialized corpora [Romer 2010] has been suggested. Incorporating these types of corpora can assist teachers in designing language analysis activities that engage students in exploring texts that are more directly relevant to their own lives. This approach allows for a more focused and meaningful analysis of language usage.

Idioms in foreign language teaching

The core of language comprises a multitude of pre-established constructions and expressions, collectively termed phraseology. Phraseology, as a fascinating field of linguistics, serves as a remarkable barometer revealing how a particular society views the world. Within the realm of phraseological units lies a treasure trove of insights into the essence of a specific culture – particularly, those idioms that are deeply intertwined with that culture. These culture-specific idioms, shaped by a multitude of external factors like climate, customs, people's habits, and the fauna in their environment, remain exclusive to certain cultures and languages, eluding discovery elsewhere.

Linguists have duly acknowledged the importance of phraseology, resulting in a burgeoning body of literature dedicated to its study. Numerous studies have explored idioms under various names, including "phraseological units", "prefabs", "formulaic expressions", "chunks", "lexical bundles", "collocation", "multi-word units", etc. [Gray & Biber 2015, p. 125].

In recent years, linguists have made a remarkable discovery regarding the profound impact of idioms and idiom-like constructions on our grasp of language. Furthermore, they have acknowledged that mastering idiomatic expressions is not just an incidental aspect but a central component of linguistic competence [Siepmann 2008, p. 186]. These findings underscore the crucial role of studying and understanding idioms in unravelling the intricacies of language.

G. Kennedy asserts that there is no one-size-fits-all method for teaching or learning idioms [Kennedy

2008, p. 23]. Second language learners typically acquire idioms through either rote memorization or inferring meanings from context. While there is a recent trend of discouraging memorization, proficient language users may still find it beneficial, as it can greatly aid comprehension and understanding, particularly when applied strategically. The effectiveness of memorization depends on the type of idiom and the learner's individual characteristics, and it can be a powerful tool in reinforcing language skills.

Scholars argue that idioms are best learned implicitly through repeated exposure in context, as encountering them frequently enhances learners' command over them. Additionally, learners often find guessing the meanings of idioms from context to be a valuable strategy [Schmitt 2008]. This suggests that a combination of contextual immersion and inference can be instrumental in the effective acquisition of idiomatic expressions.

Moreover, the meaning of a word with multiple senses (polysemous word) is often clarified by the surrounding context [Webb 2007]. Nevertheless, F. Boers and S. Lindstromberg caution against relying solely on guessing word meanings from context because learners tend to overestimate their comprehension of such words [Boers & Lindstromberg 2008]. In contrast, I. Nation asserts that repeated encounters with lexical items in various contexts are essential for effective learning, emphasizing the significance of a long-term approach [Nation 2001].

It is worth noting that the nature of context used in vocabulary acquisition research can vary significantly from one study to another. Different types of context may offer varying levels of information, leading to distinct effects on vocabulary knowledge. For instance, richer contexts are believed to have a more profound impact on vocabulary acquisition, whereas a single glossed sentence context may yield only limited gains in vocabulary knowledge [Webb 2007]. Thus, the choice of context can influence the efficacy of vocabulary learning strategies.

The introduction of corpora in language education has sparked a revolutionary shift, drawing attention to the significance of teaching multi-word expressions including idioms. This transformative approach is evident in various studies exploring the implications of the idiom and its teaching. To effectively teach idioms, digital environments, as suggested by D. Wible, prove advantageous for learners due to their dynamic, distributed, and interactive nature [Wible 2008]. However, despite these advancements, crucial aspects related to idiom learning, like the frequency of encounters and the kind of expo-

sure learners receive, remain unclear [Coxhead 2008, p. 155], presenting challenges for teachers to address.

Using COCA corpus to teach idioms in EFL classroom

The use of corpora such as the COCA emerges as a valuable avenue for understanding how these expressions function within real-world language usage. With a plethora of idiomatic expressions present in the corpus, students can gain a deeper appreciation for the interplay between language and culture. Immersing themselves in the contextual usage of idioms, learners can observe how these phrases take on nuanced meanings and reflect specific cultural nuances. The corpus not only offers a rich repository of idioms but also provides a means to examine their frequency of occurrence and variations in different registers and genres. Additionally, learners can investigate the impact of context on the interpretation and comprehension of idioms, aligning with previous insights that suggest the significance of contextual immersion in the process of language acquisition.

In this study, we aim to elucidate the algorithm and instructional activities conducive to teaching phraseology through corpus analysis. To exemplify the implementation of these techniques, we have selected a thematic group of idioms denoting youth from the Oxford Dictionary of Idioms. The idiomatic expressions under scrutiny include: 'angry young man', 'the awkward age', 'babes in the wood', bright young thing', 'a broth of a boy', 'at your mother's (or father's) knee', 'knee-high to a grasshopper', 'poor little rich girl (or boy)', 'your salad days', 'ugly duckling', 'take years off someone' [The Oxford Dictionary of Idioms 2004, p. 340].

To begin, the instructor should acquaint the students with the COCA Interface. This entails providing a comprehensive overview of the COCA interface, guiding them in the proficient access and navigation of the corpus. Essential search features and filters, instrumental in refining idiom searches, ought to be elucidated. For instance, students should be apprised of the capacity to explore idioms across diverse genres such as blogs, fiction, newspapers, and academic texts.

The teacher should actively encourage students to engage in comparative analyses of idiom usage across various genres, registers, and temporal epochs within the corpus. Using the COCA corpus as a tool, learners can thoroughly scrutinize how the frequency of different idiomatic expressions fluctuates across diverse text types. This may be presented on the example of the idioms "ugly duckling" and "babes in the wood" (Figures 1 and 2). This analytical pursuit cultivates a more profound comprehension of the dynamic evolution of language over time.

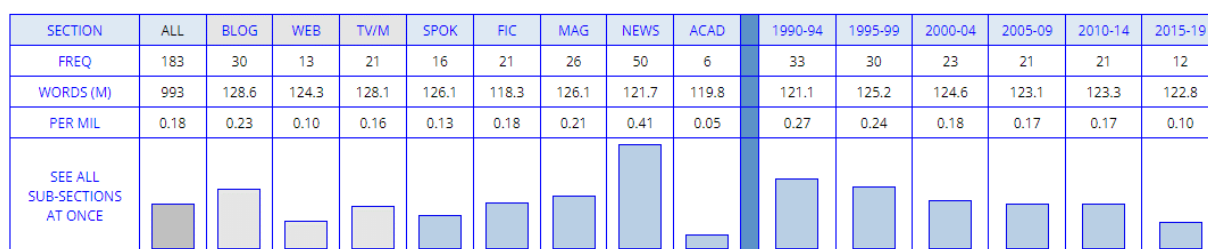| SECTION	ALL	BLOG	WEB	TV/M	SPOK	FIC	MAG	NEWS	ACAD	1990-94	1995-99	2000-04	2005-09	2010-14	2015-19
FREQ	183	30	13	21	16	21	26	50	6	33	30	23	21	21	12
WORDS (M)	993	128.6	124.3	128.1	126.1	118.3	126.1	121.7	119.8	121.1	125.2	124.6	123.1	123.3	122.8
PER MIL	0.18	0.23	0.10	0.16	0.13	0.18	0.21	0.41	0.05	0.27	0.24	0.18	0.17	0.17	0.10
SEE ALL SUB-SECTIONS AT ONCE															

Figure 1. Distribution of "ugly duckling" across genres and temporal periods (Source: Davies, 2023)

SECTION	ALL	BLOG	WEB	TV/M	SPOK	FIC	MAG	NEWS	ACAD	1990-94	1995-99	2000-04	2005-09	2010-14	2015-19
FREQ	5	1	1	1	0	0	1	1	0	0	1	0	0	2	0
WORDS (M)	993	128.6	124.3	128.1	126.1	118.3	126.1	121.7	119.8	121.1	125.2	124.6	123.1	123.3	122.8
PER MIL	0.01	0.01	0.01	0.01	0.00	0.00	0.01	0.01	0.00	0.00	0.01	0.00	0.00	0.02	0.00
SEE ALL SUB-SECTIONS AT ONCE															

Figure 2. Distribution of "babes in the wood" across genres and temporal periods (Source: Davies, 2023)

Subsequently, learners can receive guidance on the execution of keyword searches for targeted idioms. Emphasis may be placed on the omission of articles to achieve exact matches, thus ensuring heightened precision in results. As an illustration, the idiom "the awkward age" manifests in four conceivable variants within the corpus, encompassing instances with zero, indefinite, and definite articles, alongside the pronoun "that" (Figure 3).

TV	Moonlight		16? Well, going on 200. I saw a picture. Ugh. **awkward age**. You're not kidding. You know, boys that age think About
MOV	The Ice Harvest		ta give him time. He'll come around. Your leaving came at an **awkward age** for him. A clean break is best. No, a clean break
MAG	People		's fifth visit to Hogwarts is her best yet, as Harry Potter hits that **awkward age**. # PHOTO (COLOR): Brown's brainy thriller topped the bestseller
MAG	People		how old Harry Potter is in the book? Fifteen. He's at that **awkward age** -- he's too old for Michael Jackson and too young for Demi Moore
MAG	Bazaar		which they probably did. When I was 13 or 14 years old, an **awkward age** by any standard, I felt the most terrible anxiety about spending the Easter
MAG	AmHeritage		he finally lost his childhood bangs. # He was entering his adolescence, an **awkward age** for any youngster but especially so for one whose popularity
MAG	WashMonth		suppose that's why I found Didion's wry wisdom so compelling. At the **awkward age** of 17, I had bought the line that good manners, high SAT

Figure 3. The idiom "the awkward age" in four variants in the corpus (Source: Davies, 2023)

It is imperative for students to recognize that idioms can manifest in diverse parts of speech and grammatical structures. As such, the instructor's objective is to instil in them a propensity for exploring the assorted grammatical variations of idiomatic expressions. The teacher can guide students to utilize the COCA corpus to conduct searches and discern distinct renditions of the chosen idioms selecting a few idioms with multiple structural forms. For instance, the idiom "take years off" with the verbal core component may be in one of the five possible variations as presented in Figure 4. The search has been conducted capitalizing an entire word "take": "TAKE years off"

HELP	ⓘ	★	ALL FORMS 71	FREQ	
1	❶	★	TAKE YEARS OFF	40	
2	❶	★	TOOK YEARS OFF	9	
3	❶	★	TAKING YEARS OFF	9	
4	❶	★	TAKEN YEARS OFF	7	
5	❶	★	TAKES YEARS OFF	6	
			TOTAL	71	

Figure 4. The idiom "take years off" in its structural variants in the corpus (Source: Davies, 2023)

When examining idioms across languages, various transformations and adaptations become evident, reflecting the diverse ways in which language evolves and adapts to local contexts. Idioms can undergo lexical variation, where certain words within the expression are substituted with synonyms or related terms. This transformation can occur to create a similar idiomatic expression with a slightly different flavour or emphasis. Thus, students may be asked to find different versions of an idiom presented in the corpus. For instance, the lexical variations of the above given idiom may be found by entering the following search query – "[v] years off" which yields the following results presented in Figure 5.

HELP	ⓘ	★	ALL FORMS (SAMPLE: 100)	FREQ	TOTAL 103	UNIQUE 24 +
1	❶	★	TAKE YEARS OFF	32	
2	❶	★	SHAVE YEARS OFF	10	
3	❶	★	TOOK YEARS OFF	9	
4	❶	★	TAKING YEARS OFF	9	
5	❶	★	TAKEN YEARS OFF	7	
6	❶	★	TAKES YEARS OFF	6	
7	❶	★	IS YEARS OFF	6	
8	❶	★	BE YEARS OFF	5	
9	❶	★	CUT YEARS OFF	3	
10	❶	★	ARE YEARS OFF	2	
11	❶	★	WAS YEARS OFF	1	
12	❶	★	'S YEARS OFF	1	
13	❶	★	SWIPE YEARS OFF	1	
14	❶	★	SLASHED YEARS OFF	1	
15	❶	★	SLASH YEARS OFF	1	
16	❶	★	SHAVING YEARS OFF	1	
17	❶	★	PEELING YEARS OFF	1	
18	❶	★	PARE YEARS OFF	1	
19	❶	★	LIFTED YEARS OFF	1	
20	❶	★	KNOCK YEARS OFF	1	
21	❶	★	EXERCISE YEARS OFF	1	
22	❶	★	ERASE YEARS OFF	1	
23	❶	★	CUTTING YEARS OFF	1	
24	❶	★	CUTS YEARS OFF	1	

Figure 5. Lexical variations of the idiom "take years off" in the corpus (Source: Davies, 2023)

The analysis shows the idiom appears in various lexical variations substituting the verb "take" with the lexemes "shave", "cut", "swipe", slash", "peel", "knock", "erase", etc. Subsequently, students are prompted to compare the meanings and usages of these diverse iterations across various genres. This educational exercise develops students' ability to identify variations and nuances within idiomatic expressions, fostering a more comprehensive understanding of their usage and significance.

Another activity involves conducting a comparative analysis between a specific idiom and the collocations formed by its constituent lexical components. This activity aims to deepen students' understanding of idiomatic expressions and their underlying linguistic structures. The search query to consider the idiom "salad days" may be formulated as "salad [n]" providing the following results (Figure 6):

HELP	ⓘ	★	ALL FORMS (SAMPLE): 100 200 500	FREQ	TOTAL 3,614 | UNIQUE 553 ◆
1	ⓘ	★	SALAD BAR	423	
2	ⓘ	★	SALAD GREENS	419	
3	ⓘ	★	SALAD BOWL	293	
4	ⓘ	★	SALAD DRESSINGS	249	
5	ⓘ	★	SALAD DAYS	139	
6	ⓘ	★	SALAD SANDWICH	132	
7	ⓘ	★	SALAD PLATES	100	
8	ⓘ	★	SALAD BARS	98	
9	ⓘ	★	SALAD PREP	86	
10	ⓘ	★	SALAD OIL	81	
11	ⓘ	★	SALAD PLATE	65	
12	ⓘ	★	SALAD SANDWICHES	65	
13	ⓘ	★	SALAD SPINNER	64	
14	ⓘ	★	SALAD INGREDIENTS	60	
15	ⓘ	★	SALAD MIX	60	
16	ⓘ	★	SALAD FORK	57	
17	ⓘ	★	SALAD DRESSING	38	
18	ⓘ	★	SALAD BOWLS	29	
19	ⓘ	★	SALAD COURSE	27	

**Figure 6. The idiom "salad days" vs. collocations with the lexeme "salad"
in the corpus (Source: Davies, 2023)**

Examining the collocational patterns associated with the individual words composing the idiom, learners can study the semantic relationships and contextual usage of these components apart from the frequency of their usage. This exploration fosters a more nuanced comprehension of how idioms and collocations function within the broader framework of phraseology and enhances students' ability to employ them effectively in their language usage.

Significant emphasis is attributed to the criticality of contextual analysis in comprehending idiomatic expressions. The instructor should engage in discourse with students, focusing on the value of examining the contextual environment in which idioms are situated. Discerning the co-occurring words and sentences surrounding an idiom, invaluable insights can be obtained regarding its intended meaning and usage. Such discussions encourage learners to recognize the subtle intricacies inherent in the application of idiomatic expressions within real-world language contexts.

An effective instructional strategy involves providing students with a curated selection of idiomatic expressions and tasking them to compose dialogues or short stories incorporating these idioms. In this creative exercise, the COCA corpus serves as a valuable reference tool, enabling students to verify the correct usage and appropriate contexts of the idiomatic expressions they employ. In this way, learners can refine their ability to effectively integrate idioms into authentic language contexts, bolstering their language proficiency and creative writing skills.

Beyond the classroom, the knowledge gained from studying idioms in the COCA corpus holds practical applications in real-life scenarios, particularly in understanding colloquial speech and informal writing. Familiarity with idiomatic expressions gleaned from the corpus enables individuals to grasp the subtleties and nuances inherent in informal language use, facilitating more effective communication in social settings. Moreover, when encountered in various genres, idioms drawn from the corpus provide students with deeper insights into the cultural and societal underpinnings of the language. This heightened understanding of idiomatic language not only enriches language comprehension but also enhances the ability to engage effectively in informal conversations, contributing to an overall sense of language fluency.

Conclusions. Corpus linguistics has brought about a fundamental shift in our comprehension of language usage, consequently influencing how we approach language education. In the coming decades, corpus-based findings are likely to continue shaping the development of language teaching materials. The employment of corpora in the classroom is also expected to increase, with teachers and students gaining greater access to online corpora and text analysis tools. Rather than questioning the relevance of corpus linguistics to language teaching, the focus now lies in determining the most effective ways for teachers and students to use the available corpus-based resources. Through further exploration of corpus usage in diverse classroom contexts, our understanding of the various applications of corpus linguistics in language pedagogy will undoubtedly be refined.

Engaging with the COCA corpus encourages students to deal with the complexities of idioms and idiom-like constructions, honing their phraseological com-

petence and gaining a more profound mastery of linguistic competence. The repeated exposure to idioms within diverse contexts facilitated by the corpus fosters a long-term approach to learning, leading to a more solid understanding and fluency in using idioms. While a corpus is not a failsafe method of learning idioms, it certainly offers a practical and comprehensive means of analysing and understanding these linguistic expressions. Using a corpus, learners can improve their idiom learning strategies by going beyond simple memorization and embracing the important aspect of contextual understanding.

In addition, employing the COCA corpus in the study of idioms enriches the pedagogical toolkit, offering a nuanced and data-driven approach to language learning. This corpus exploration not only allows students to explore culture-specific idioms but also provides a broader perspective by comparing idiom usage across diverse genres and historical periods. Fostering a collaborative learning environment, where students can share their findings and interpretations, the journey of learning idioms becomes engaging and dynamic. With a vast array of idioms to explore and a profound understanding of how context shapes their meanings, students can attain a higher level of idiomatic competence, making significant strides towards enhancing their overall language proficiency.

References

1. Aston G. (2001) Learning with corpora: An overview. *Learning with corpora* . Ed. G. Aston. Houston, TX: Athelstan. P. 7–45 [in English].

2. Bernardini S. (2001) "Spoilt for choice": A learner explores general language corpora. *Learning with corpora*/ed. G. Aston. Houston, TX: Athelstan. Houston, TX: Athelstan. P. 220–49 [in English].

3. Biber D., & Reppen R. (2002) What does frequency have to do with grammar teaching? *Studies in Second Language Acquisition*, 24. P. 199–208 [in English].

4. Boers F. & Lindstromberg S. (2008) From empirical findings to pedagogical practice. *Cognitive Linguistic Approaches to Teaching Vocabulary and Phraseology*. Ed. Boers, F. and Lindstromberg, S. Berlin: Mouton de Gruyter. P. 375–393 [in English].

5. Cobb T. (2007) Computing the vocabulary demands of L2 reading. *Language Learning and Technology*, 11. P. 38–64 [in English].

6. Conrad S. (2000) Will corpus linguistics revolutionize grammar teaching in the 21st century? *TESOL Quarterly*, 34. P. 548–560 [in English].

7. Coxhead A. (2008) Phraseology and English for academic purposes: Challenges and opportunities. *Phraseology in Foreign Language Learning and Teaching*. Ed. F. Meunier & S. Granger. Amsterdam: John Benjamins. P. 149–161 [in English].

8. Davies M. (2023) The corpus of contemporary American English: 520 million words, 1990-present. Available from https://www.english-corpora.org/coca [in English].

9. Flowerdew L. (2009) Applying corpus linguistics to pedagogy: A critical evaluation. *International Journal of Corpus Linguistics*, 14. P. 393–417 [in English].

10. Granger S. (2002) A bird's-eye view of learner corpus research. *Computer learner corpora, second language acquisition and foreign language teaching*. Ed. S. Granger, J. Hung, &: S. Petch-Tyson. Philadelphia, PA: John Benjamins. P. 3–33 [in English].

11. Gray B. & Biber D. (2015) Phraseology. *Cambridge Handbook of Corpus Linguistics*. Ed. Biber & R. Reppen. Cambridge: Cambridge University Press. P. 125–145 [in English].

12. Hunston S. (2002) Corpora in applied linguistics. Cambridge, England: Cambridge University Press [in English].

13. Johns T. (2000) Data-driven learning: The perpetual challenge. *Teaching and learning by doing corpus analysis: Proceedings of the Fourth International Conference on Teaching and Language Corpora*. Ed. B. Kettemann & G. Marko. New York, NY: Rodopi. P. 107–118 [in English].

14. Kennedy G. (2008) Phraseology and language pedagogy: Semantic preferences associated with English verbs in the British National Corpus. *Phraseology in Foreign Language Learning*. Ed. Meunier, F. and Granger, S. Amsterdam/Philadelphia: John Benjamins Publishing Company, P. 21–41 [in English].

15. McCarthy M. (2008) Accessing and interpreting corpus information in the language education context. *Language Teaching*, 41. P. 563–74 [in English].

16. Nation I.S.P. (2001) Learning Vocabulary in Another Language. Cambridge: Cambridge University Press [in English].

17. Romer U. (2010) Using general and specialized corpora in English language teaching: Past, present and future. *Corpus-based approaches to English language teaching*. Ed. M. C. Campoy-Cubillo, M. Luisa Gea-Valor, & B. Belles-Fortuno. London, England: Continuum. P. 18–35 [in English].

18. Siepmann D. (2008) Phraseology in learners' dictionaries: What, where and how? *Phraseology in Foreign Language Learning*. Ed. Meunier, F. and Granger. Amsterdam/Philadelphia: John Benjamins Publishing Company. P. 185–202 [in English].

19. Sinclair J. (2005) Reading concordances. London, England: Longman [in English].

20. Schmitt N. (2008) Review article: Instructed second language vocabulary learning. *Language teaching research*. Vol. 12 (3). P. 329–363 [in English].

21. The Oxford Dictionary of Idioms (2004). Ed. Judith Siefring. Oxford University Press [in English].

22. Webb S. (2007) Learning word pairs and glossed sentences: The effects of a single context on vocabulary knowledge. *Language teaching research*. Vol. 11 (1). P. 63–81 [in English].

23. Wible D. (2008) Multiword expressions and the digital turn. *Phraseology in Foreign Language Learning and Teaching*. Ed. F. Meunier & S. Granger. Amsterdam: John Benjamins. P. 163–184 [in English].

ВИКОРИСТАННЯ КОРПУСНОГО ПІДХОДУ У ВИВЧЕННІ АНГЛІЙСЬКИХ ІДІОМ

Анотація. Ідіоматичні вислови як відображення культури, які глибоко вкорінені в мові, дають цінне уявлення про соціальні цінності, звичаї та вірування. У цій статті досліджується важлива роль, яку ідіоми відіграють у викладанні іноземної мови, з особливим фокусом на такому потужному інструменті, як Корпус сучасної американської англійської мови (COCA), для покращення рівня володіння мовою та розуміння культури.

Дослідження, присвячене використанню корпусної лінгвістики у навчанні ідіом на заняттях з англійської мови, є надзвичайно актуальним, оскільки воно вивчає методи навчання мови, що сприяють збагаченню мовних навичок та підвищенню міжкультурної обізнаності, даючи можливість викладачам приймати обґрунтовані рішення для ефективного викладання.

Мета цього дослідження – доповнити наявну літературу з досліджуваної проблеми та представити алгоритм та навчальні вправи для викладання ідіом на основі корпусного аналізу, пропонуючи цінну інформацію для викладачів іноземних мов, які прагнуть оптимізувати процес навчання своїх студентів.

Методологія охоплює керівництво для студентів у доступі до корпусу та навігації по ньому, проведення цілеспрямованого пошуку ідіом, вивчення їхніх варіацій та контекстуального використання, а також їхнього залучення до навчання з метою покращення розуміння та використання ідіоматичних висловів.

Ознайомлюючи учнів із контекстуальними ідіомами в різних жанрах, регістрах та історичних періодах, викладачі можуть виховати в них глибоке розуміння динамічного розвитку мови в часі. Підкреслюється важливість контекстуального аналізу, який допоможе навчити студентів розрізняти найтонші нюанси та відтінки значення у вживанні ідіом. Залучення студентів до творчих вправ, зокрема таких, як складання діалогів та історій, може сприяти практичному застосуванню ідіом і посилити їхню інтеграцію в автентичний мовний контекст. У цій статті підкреслюється важливість виваженого використання корпусу як важливого ресурсу на заняттях з англійської мови викладачами, які прагнуть допомогти своїм студентам у вивченні багатого розмаїття ідіом та глибшому розумінні мови та культури.

Ключові слова: ідіома, корпус, корпусна лінгвістика, Корпус сучасної американської англійської мови, навчання англійської мови як іноземної.

Maryana Tomenchuk – Candidate of Philology, Associate Professor of the Applied Linguistics Department, Uzhhorod National University, Uzhhorod, Ukraine; marianna.tomenchuk@uzhnu.edu.ua; https://orcid.org/0000-0002-2036-4616

Мар'яна Томенчук – кандидат філологічних наук, доцент кафедри прикладної лінгвістики Ужгородського національного університету, Ужгород, Україна; marianna.tomenchuk@uzhnu.edu.ua; https://orcid.org/0000-0002-2036-4616