

Corpus-based Analysis of Lexicosemantic Behaviour of Nervous System Diseases Names

Marta Lacková

Department of English Language and Literature
Faculty of Humanities, University of Žilina
Žilina, Slovakia

Abstract:

Terminology represents a significant factor in healthcare communication between specialists and patients. The present paper deals with the lexicosemantic characteristics of multi-word lexical units *multiple sclerosis*, *amyotrophic lateral sclerosis*, *Parkinson's disease*, *Alzheimer's disease*, *epidural abscess*, *Huntington's chorea*, and *carpal tunnel syndrome*. The research questions focus on collocations, word combinations, and concordances in which they regularly appear; the first lemma to the left and the first lemma to the right from the studied lexeme demonstrate the principal positions of our interest. Simultaneously, the researcher considers their semantic restrictions, semantic prosody, and grammatical relations that influence their lexical features. The whole linguistic material is investigated in the framework of the text corpus English Web 2015 (enTenTen15) with the help of the search tool Sketch Engine. To begin the research, the frequencies of these lexical units are elaborated. The researcher also looks into the morphological classification of the studied words as these two factors affect them from the lexicological perspective. The research outcomes confirm that the nervous disease names appear in a wide range of structures, and they considerably contribute to successful communication in the medical surroundings. Moreover, the results indicate that the lexicosemantic behaviour of the terms reflects extralinguistic factors (psychological, social) of individual communication acts. The phenomenon is to be further examined and interpreted within the corpus analyses of other chosen lexical units, not only from the field of medicine. Eventually, the researcher outlines possible pedagogical implications of the research results in the process of teaching English.

Keywords: corpus-based analysis, collocation, concordance, lemma, lexicosemantics, meaning, nervous system disease names

Cite as: Lacková, M. (2020). Corpus-based Analysis of Lexicosemantic Behaviour of Nervous System Diseases Names. *Arab World English Journal*, 11 (4) 50-65.

DOI: <https://dx.doi.org/10.24093/awej/vol11no4.4>

Introduction

The textual or linguistic corpus is defined as an extensive collection of linguistic information and texts elaborated with the help of computers, which is applicable for research in linguistics (Baker, Hardie & McEnery, 2006). Sinclair (1991, p. 14) provided a similar definition of corpus; he defines a corpus as “a collection of naturally occurring language texts, chosen to characterize a state or variety of a language”. In recent decades, research papers applying various corpus linguistics techniques to the study of diverse linguistic phenomena have been increasing. As one of the most significant developments in vocabulary studies, it delivers an empirical basis for determining vocabulary behaviour instead of relying on tradition or intuition. Consequently, corpora have transformed how we research vocabulary systems (Schmitt, 2010; Friginal, 2017).

Corpus linguistics is a branch of linguistics that investigates languages based on discourse. Regarding this, corpus linguists have to find a suitable sample of the discourse to work with; the sample is called the corpus (Halliday, Čermáková, Teubert & Yallop, 2004, p. 100). Corpus linguistics, corpus tools, and corpus evidence offer a broader and more delicate perspective into the language in use, to understand how language works in specific contexts (Leláková, 2018).

Moving to the specific area of semantics aspects investigation, the researcher considers words in context. Therefore, corpus concordances enable us to realize how much the meanings of words derive from context; there raises a controversial question of whether words have an independent sense at all (Moon, 2006). One possible response is that corpora show us specific contexts in which a word appears, and it is our knowledge of the word that relates it to its associations with other words (McEnery & Hardie, 2006).

Primarily, the lexical units *multiple sclerosis*, *amyotrophic lateral sclerosis*, *Parkinson's disease*, *Alzheimer's disease*, *epidural abscess*, *Huntington's chorea*, and *carpal tunnel syndrome* have been chosen as a research sample for the purpose to specify the structures in which they emerge, and to deduce the lexicosemantic characteristics arising from them. Based on the theoretical background, three research questions have been formulated:

1. What are the most typical collocations in which the nervous system disease names appear?
2. What semantic qualities reflect the adjectives placed before the nervous system disease names?
3. What is the capability of nervous system disease names to be predetermined by formal expressive means of the morphological category of determination?

The principal objective of the research is to study collocations, word combinations, and concordances in which the lexical units *multiple sclerosis*, *amyotrophic lateral sclerosis*, *Parkinson's disease*, *Alzheimer's disease*, *epidural abscess*, *Huntington's chorea*, and *carpal tunnel syndrome* regularly appear. The predominant analysed position is the first lemma to the left and the first lemma to the right from the studied lexical unit. The following aim is to outline, summarize, and exemplify their typical lexicosemantic features. Simultaneously, the research intention is to draw general conclusions about the functioning of nervous system disease names in the textual corpus English Web 2015 (enTenTen15).

Literature Review

Language use is of evident practical importance to healthcare professionals, healthcare students who communicate medical ideas regularly, and non-professionals, who constitute their experience of health and illness. Healthcare communication and its various aspects have become a significant theme of research within Applied Linguistics. Previous studies on health communication have contributed significantly to the analysis of its sociological factors; simultaneously, it is necessary to highlight their close focus on language in use to point up the leading role of language in the practice of health care and medicine (Svenja, Brown, Carter, Crawford & Sahota, 2004). The processes of globalisation are reflected in the change of lexical composition of the language, as new forms of professional relations between people are appearing (Semeniuk, Leleka, Moskalenko, 2020).

Furthermore, several health communication studies have applied corpus methodologies. However, they rarely appear in comparison with other linguistic themes from different branches, such as lexicography, lexicology, morphology (Fellbaum, 2007; Gavioli, 2005; Kadorová & Ondráčková., 2005; Jones & Waller, 2015). More recently, Skelton, Wearn & Hobbs (2002) demonstrated the advantages of integrating qualitative with quantitative approaches in the framework of corpus linguistics. They claim that it is only qualitative methods, such as concordance lines with subsequently extended stretches of text that can explain the available patterns of health language functioning. By way of contrast, quantitative methods identify general patterns that, exist in a complex context describable only partially quantitatively (Atkins & Harvey, 2006, Flowerdew, 2004). The recent approaches emphasize the selection of optimal significance tests, and estimate the similarity of distribution patterns (Wallis, 2020).

Regarding other linguistic research methods, a comparative approach was adopted by (Svenja et al., 2004); they compare a corpus of health professionals' language with a corpus of general spoken English and identify a set of keywords that appear with greater frequency in professional communication. These methodological procedures enabled the authors to characterize specific communication while identifying its prevailing tendencies.

From the sociolinguistic perspective, McEnery & Hardie (2006, p. 117) emphasize that "it is becoming more and more common for sociolinguistically sampled data to be compiled and described in a corpus-like manner". So far, the study of nervous system illness and disease names in the corpus environment has not been a topic of considerable interest; out of these reasons, the researcher strives to uncover this linguistically challenging issue. The researcher's aim of investigating these lexical units in a corpus framework has been motivated by the claim that the linguistic material would be best described on examples that appear in natural discourse. Corpus linguistic methodologies will be adopted to research in this field.

Data and Methodology

For the needs of the research, the researcher has worked with the English monolingual synchronic annotated textual corpus English Web 2015 (enTenTen15). As a corpus-based paper, it aims to identify recurrent patterns in which the keywords occur to examine their discourse functions, lexical and semantic peculiarities, and sociological factors in the background.

To search for, sort, and classify the chosen lexical units *multiple sclerosis*, *amyotrophic lateral sclerosis*, *Parkinson's disease*, *Alzheimer's disease*, *epidural abscess*, *Huntington's chorea*, and *carpal tunnel syndrome* in the corpus, the researcher will utilize methods of corpus linguistics together with statistical methods interconnected in the unique tool for searching in corpus Sketch Engine. It is essential to emphasize that, on the one hand, the corpora allow the scholar to see the language in the context of discourse. On the other hand, the language samples provided by the corpora do not immediately display specific linguistic phenomena in the most general way. On the grounds of this, the researcher will also apply methods of generalization and classification of linguistic phenomena.

The researcher takes the computerized corpora and computational tools as a starting point for a qualitative analysis, enabling her to study the immediate discourse in which the keywords occur together with broader psychological and social contexts in which they find their positions. The principal criterion is that the researcher selects only statistically significant words for our analyses; the researcher excludes lemmas that display frequency pre million < 0.01 .

Overall, the choice of nouns for a more in-depth analysis has been reasonably motivated by their frequency. However, scientists understand frequency as a relative phenomenon; different corpora result in other frequency lists. According to Mahlberg (2005), single words in frequency lists give information about a text's content; nevertheless, their position in the list does not tell us much about meaning. As Stubbs (2001) points out: "The computer findings (frequencies, comparisons or graphs) are not the interpretation of the meaning of the text, but the presentation of some of its formal features". These standard features have to be linked to contextual features to uncover the meaning of the studied words. Finally, some of the analysed lexical units belong to the group of compound nouns. Scholars do not view the concept of the unit of meaning as the criterion for fixed expressions as arbitrary (Halliday et al., 2004).

Another important point about the selection of nouns for this study is the fact the number of nervous system disease names that the researcher could handle here is limited. The aim has been merely to work with a manageable amount of information that is to be presented within this paper. To build a complex picture of the semantic behaviour of names of nervous system diseases, the researcher has selected nouns/compound nouns in the three positions of the frequency scale:

- lexical units with high frequency: *multiple sclerosis*, *amyotrophic lateral sclerosis*, *Parkinson's disease*, *Alzheimer's disease*;
- lexical units with frequency in the middle of the scale: *carpal tunnel syndrome*;
- lexical units with low frequency: *epidural abscess*, *Huntington's chorea*.

Table 1. *Number of concordances of the studied lexical units and their frequency per million in English Web 2015 (enTenTen15)*

Nervous system disease	Number of concordances	Frequency per million
Alzheimer's disease	66 368	3.61
Parkinson's disease	44 109	2.4
multiple sclerosis	43 258	2.35
amyotrophic lateral sclerosis	5322	0.29
carpal tunnel syndrome	4127	0.22
epidural abscess	259	0.01
Huntington's chorea	169	0.01

Results

The first phase of the research is represented by observing collocations, combinations, and concordances of the studied lexical units in the corpus English Web 2015 (enTenTen15). From a methodological procedure point of view, it is more convenient to treat a phrase as a collocation than to describe it as the co-occurrence of two single words. This section outlines, summarizes, and exemplifies the qualitative analysis results together with the semantic implications emerging from them.

Alzheimer's Disease

The corpus evidence manifests the occurrence of compound noun *Alzheimer's disease* in a wide range of structures. Its word-class patterns are apparent: the compound noun prevails with determiners; different modifiers ranging from adjectives to nominal and verbal prepositional phrases modify the lexical unit.

Adjectives (marginally of domestic origin) that in the English Web 2015 (enTenTen15) Corpus precede the compound noun *Alzheimer's disease* are classified from the point of view of their semantic signs in the following way:

- adjectives denoting temporal aspects: *sporadic, developing, onset, late-onset, early-onset, advancing*;
- adjectives denoting the subjective attitude (the most extensive subgroup): *mild, mild-to-moderate, moderate*, while *severe Alzheimer's disease* is the prevailing combination in this context: the power of emotive words has been recognized and proved;
- adjectives connecting *Alzheimer's disease* with the psychic state of the patient: *dementia, pre-dementia*.

In contrast, there was no increased frequency in *early-onset sporadic Alzheimer's disease* or patients with vascular dementia. (bio2rdf.org)

The evidence obtained from the corpus implies that verbs that expand the elements of meaning of the compound noun *Alzheimer's disease* appear in more semantically relevant collocations than regular dictionaries provide us with: *treat, develop, prevent, moderate, diagnose, detect, cause, fight, combat*. What is more, the verb *face* displays metaphorical connotations within six concordances:

The clinic provides diagnosis, assessment, and research for thousands of individuals who *face* Alzheimer's *disease* and other memory disorders. (mind.uci.edu)

The most frequent lexical units in the position the first lemma to the left from the compound noun *Alzheimer's disease* are prepositions creating prepositional phrases; they are the descriptions of diverse external conditions connected with the treatment of the disease: the preposition *of*: (*the treatment of*); the preposition *with*: (*affected with*); the preposition *in*: (*progress in*), the preposition *for*: (*immunotherapies for*); the preposition *about*: (*think about*).

The nouns often associate with emotions, and these are predominantly negative ones. The researcher considers about one-third of the concordances to have a neutral tone, and only one-tenth of them are of a positive manner. The contribution of these nouns to the communicative message becomes clear thanks to the specific context in which they appear.

Seen from the perspective of nouns semantics and hyponymic relations, the compound noun *Alzheimer's disease* has the potential to be combined with various conditions; but still there emerge several subgroups which unite meanings further specifying its features or signs: *cataracts, diabetes, obesity, inflammation, hypertension, atherosclerosis, Parkinson's disease, Alzheimer's disease, nervous system deterioration, muscle breakdown, chronic fatigue syndrome*. The specific example below highlights the point mentioned above; at the same time, it indicates a possible connection between *Alzheimer's disease* and the mental health of the patient:

Memory problems. Stress. Confusion. Bizarre behaviour. Depression. Anxiety. Caregiver burnout. These challenges all too often occur in the territory of *Alzheimer's Disease* and related dementias (ADRD). (*healthpassion.info*)

Another way of examining the compound noun *Alzheimer's* is to observe its capability to be predetermined by the definite or indefinite articles. The definite article is utilized in 1430 concordances as the first lemma to the left; it fulfils two primary functions in this position:

- the cataphoric reference (we learn more about the disease later from the context): *to find the cure for the Alzheimer's disease that had claimed his grandmother*;
- *to modify a noun following the disease's name, not the disease itself (part of a proper name): The Alzheimer's Disease Center, the Alzheimer's Disease Neuroimaging Initiative, the Alzheimer's Disease Fund.*

The indefinite article appears in 248 concordances with *Alzheimer's disease*; it executes solely one function – the indefinite article with the descriptive role; the noun going after *Alzheimer's disease* is specified by the health condition: *an Alzheimer's disease patient, treatment, blood test*.

Parkinson's Disease

Studying the significant collocations in which the compound noun *Parkinson's disease* displays its potential, the researcher commences with adjectives since their use in the position first lemma to the left from the compound noun *Parkinson's disease* is particularly noteworthy; on the one hand, they express meanings that the speakers of the English language naturally utilize in connection with health conditions: *advancing, recessive, familial, idiopathic, early, sporadic, inherited, untreated Parkinson's disease*. On the other hand, there are present several emotionally coloured adjectives as in the example below: *severe, moderate, promising Parkinson's disease* and marginally adjectives associated with the mental health of the patient: *debilitating Parkinson's disease*.

Concerning the verbs that express procedural meanings in the surroundings of the compound noun *Parkinson's disease*, they display wide semantic variations on the scale from positive to negative connotations: *treat, battle, cure, diagnose, study, prevent Parkinson's disease* on one end of the plate and *develop, cause Parkinson's disease* on its other end.

The most frequent nouns occurring in the hyponymic relation with the compound noun *Parkinson's disease* include the subsequent structures: *stroke, cancer, heart disease, blindness, diabetes*. They possess the capacity to denote more health conditions existing together with *Parkinson's disease*, with which patients have to struggle.

Formal expressive means of determination are applied with the compound noun *Parkinson's disease* in the English Web 2015 (enTenTen15), hence to a lesser degree than with the compound noun *multiple sclerosis*. The application of possessive pronouns *her, their, my, our, his* in front of the health condition's name demonstrates its direct influence on the patient's life. Furthermore, the researcher spots collocations with the quantifier *all*. They modify a noun that follows the analysed lexical unit: *all Parkinson's disease patients, all Parkinson's disease sufferers*. The determiners belong to the linguistic signs signifying a relationship between the words and external discourse reality. In contrast, being of peripheral importance within the English Web 2015 (enTenTen15), the indefinite article *a* is applied in 243 concordances. The researcher again acknowledges it as having a descriptive role not for the compound noun *Parkinson's disease* itself, but for a noun placed after it: *Parkinson's disease model, patient, target*.

The structure a noun/verb + the preposition *about* + *Parkinson's disease* possesses the capacity to indicate several social and psychological factors in the studied discourse – we make sense of the social experience of health through it; the examples are literal, but such uses also imply psychological states of the members of communication acts: *learn the basics about, information about, learn more about, articles about, talking about, awareness about, myths about, facts about, questions about, a new song about, tweet about, presentations about*.

There's a new song about Parkinson's Disease called "Blame it on the Parkinson's". (wilkins-pf.org)

The structure of a noun/verb + the preposition *with* + *Parkinson's disease* is still very much associated with people connected with the disease both directly and indirectly: *diagnosed with,*

patients with, individuals with, persons with, people with, adults with, subjects with, sufferers with, struggling with, affected with.

Most *patients with Parkinson's disease* (PD) show unilateral motor impairment. (fapesp.br)

Multiple Sclerosis

Within the adjectival collocations (we search for the first adjective to the left from the studied lexical unit) in which the compound noun *multiple sclerosis* appears, the adjective *progressive* is the most frequent one that emerges with it identifying its most predictable feature. Further, the compound noun *multiple sclerosis* is preceded by the adjectives of both: native: (*battling, treating*) and Latin or Greek origin: (*relapsing, remitting, chronic-progressive, secondary-progressive, paediatric multiple sclerosis*).

In the example below, the meaning of multiple sclerosis is specified via the application of two following adjectives:

The report deals with values of cell number and total protein in *intermittent and chronic-progressive multiple sclerosis* taking into account age and sex. (scda.org.sg)

The spread of the compound noun *multiple sclerosis* in various registers of the English language is also illustrated by formal expressive means of category of determination applicable with it. *Multiple sclerosis* belongs to few lexical units within our research material that are further determined by possessive pronouns *your, their*, by the quantifier *all*, and the cardinal numeral *first*.

To make sure you're adequately equipped for *your multiple sclerosis*, it's crucial to develop a comfortable relationship with your physician and care team, Beaver said. (healthpassion.info)

As for the nouns that become known around the compound noun *multiple sclerosis* in the equal syntactical position as hyponyms, they are semantically related to further health conditions that are associated with it, and they create hyponymic chains of nervous and other body systems diseases: *rheumatism, hepatitis C, tuberculosis, rheumatoid arthritis, psoriasis, lupus erythematosus, cerebral palsy, spina bifida*. The example below represents part of a longer stretch of text that we analyse in terms of patterns as corpus linguistic conclusions presuppose that meaning is observable through designs (Mahlberg, 2005, p. 162).

His clinical interests and research centre on clinical and biological aspects of multiple sclerosis and related inflammatory diseases and on developing adult stem cell repair treatments for multiple sclerosis and other neurodegenerative conditions. (adultstemcellconference.org)

Significant collocations of the compound noun *multiple sclerosis* are observable when it becomes part of a phrase modifying another noun (we study the first lemma to the right from the studied lexical unit): *a potential multiple sclerosis treatment, sufferer, symptoms, therapy, diagnosis*. The collocations might be in concord with already established notions in dictionaries or display original word combinations: *multiple sclerosis is not a disease but a condition*. From a sociolinguistic point of view, they are noteworthy for the understanding of humans and their

behaviour and institutions. In addition to this, a limited number of these nouns have the potential to appear also in front of the analysed compound noun: *therapy of, diagnosis of multiple sclerosis, treatment of multiple sclerosis*. Their possible double position demonstrates their extralinguistic importance manifested in the communication situations between health specialists and patients.

The compound noun *multiple sclerosis* is part of a complex network of textual relations. The examples of the continuity function are also evidence of the textual complexity on which the creation of meaning becomes dependent. To illustrate this, the cases when the prepositions *in* or *on* are relevant demonstrate the scientists' interest in the disease research: *studies in multiple sclerosis, research in multiple sclerosis, a trusted source of information on multiple sclerosis*.

With regards to the preposition *with* (5589 concordances) (0.3 per million) that is the most frequent lemma to the left from *multiple sclerosis*, it is manageable to decode elements of meaning that communicate the considerable impact of the health condition on the patients' lives: *people living with diagnosed with, battle with, struggle with, experiences with multiple sclerosis*.

Furthermore, we frequently find the preposition *of* in the genitive phrase modifiers: *features of, aspects of, the prognosis of, a sufferer of, signs of, manifestations of, stages of, forms of, a victim of*. Additionally, there exist combinations with the compound noun *multiple sclerosis* directly expressing the variety of the health condition, as illustrated in the following examples: *subsets of, variant of, subgroups of, classification of*.

Amyotrophic Lateral Sclerosis

When investigating and understanding the language both doctors and patients use concerning the compound noun *amyotrophic lateral sclerosis*, we observe only a restricted number of lexical units in the first lemma to the left from it. This particular piece of information closely connects its terminological nature and etymological properties.

To identify statistically significant adjectives that provide a useful insight into the semantic behaviour of the compound noun, we spot only these in the English Web 2015 (enTenTen15) with the frequency higher than 0.1 per million lemmas: *sporadic, familial, inherited, juvenile advanced, developing amyotrophic lateral sclerosis*. Moreover, one can find the verbs *defeat, battle, treat, cure, combat, cause amyotrophic lateral sclerosis* plentifully in the same position.

Could an overeager response to viral infections predispose certain people to *juvenile amyotrophic lateral sclerosis*? (researchals.org)

When searching for additional sociolinguistic aspects that are covered by the structure a noun + a preposition + *amyotrophic lateral sclerosis*, the subsequent dominant communication elements are classified:

- information spread about the disease: *schooling about, learning more about, know about, scientific knowledge about;*
- consequences (often lethal) of the disease: *suffer from, mortality from, die from.*

The excess *mortality from amyotrophic lateral sclerosis* seems to be associated with above-average exposure to electromagnetic fields and may be due to repeated episodes with electric shocks. (emfs.info)

The position of the preposition *with* in the structure mentioned above deserves particular regard; it enables to name those affected by the health condition. Among the most frequent collocates with the preposition are the following words: *patients with, subjects with, individuals with, people with amyotrophic lateral sclerosis*.

In the specific case of *amyotrophic lateral sclerosis*, the determinate power of the definite and indefinite articles is suppressed to a minimal extent. The definite article (79 concordances) is suggested on the whole in proper name structures: *The Amyotrophic Lateral Sclerosis Association*. Scholars recommend the indefinite article (13 concordances) when the member of the communication act describes the noun in more detail: *an amyotrophic lateral sclerosis research project, an amyotrophic lateral sclerosis clinical trial*.

Carpal Tunnel Syndrome

Analysing the first lemma to the left, the compound noun carpal tunnel syndrome appears to be semantically further modified by several different word classes; the adjectives in the position are twofold:

- *denoting subjective qualities (prevailing): developing, severe, conquering, mild, moderate, recurrent; it is noticeable that there exist even cases when they contradict each other;*
- *denoting objective qualities: idiopathic, bilateral, work-related carpal tunnel syndrome.*

Aetiology of work-related carpal tunnel syndrome: the role of lumbrical muscles and tool size on carpal tunnel pressures. (mayo.edu)

Moving to the word class of verbs, they are the first lemma to the left from the compound noun *carpal tunnel syndrome* only of a limited number: *include, prevent, develop, avoid, relieve, diagnose, identify, eliminate, treat carpal tunnel syndrome*. As it is evident from this enumeration, most of them express meanings to help people suffering from the condition. The only exception out of this model is the verb *have*. *It appears to be the most frequent verb with general meaning in this setting.*

The ergonomic design helps prevent carpal tunnel syndrome while working at a keyboard. (odecompany.net)

In the English Web 2015 (enTenTen15) Corpus, there were recorded numerous nominal structures with prepositions *with* and *of* which add extra information to the compound noun *carpal tunnel syndrome*; it is evident that they facilitate associating different senses with different patterns or structures. By way of example, practices in the first subgroup express a reference point for a person affected by the health condition: *patients with, individuals with, people with; symptoms of, treatment of, studies of, the study of*.

Current projects include a large, prospective *study of carpal tunnel syndrome* and other upper extremity disorders. (wustl.edu)

The references with *carpal tunnel syndrome* found in the English Web 2015 (enTenTen15) Corpus mirror the language users' understanding of reality, their attitudes, and either objective or subjective evaluations:

- the generic use of the indefinite article (23 concordances) demonstrating any representative member of the class;

A clinical investigation can usually confirm or exclude a carpal tunnel syndrome or an ulnar nerve entrapment at the elbow or wrist level. Still, sometimes a neurophysiologic investigation is necessary to confirm the diagnosis. (vibrosense.eu)

- the indefinite article with a descriptive role (6 concordances) providing more precise information about the noun; *a severe carpal tunnel syndrome*;
- the definite article with the direct anaphoric reference (76 concordances) with the noun mentioned earlier in the context:

Epidural Abscess

The compound noun *epidural abscess* appears in discourses abundant in terminology that are not intended for the general public as the applied vocabulary makes the understanding more complicated. Mostly adjectives of Latin or Greek origin form the field of medicine *spinal, intracranial, lumbar, cervical, spontaneous, anterior, paraspinal* (25) are localized in the first lemma to the left from the compound noun *epidural abscess*. The only exceptions out of this model are the adjectives *chronic epidural abscess, untreated epidural abscess, and the spontaneous epidural abscess* that are more general.

On the same grounds, we distinguish lexical units of Greek and Latin origin in the structure: a noun + the preposition *with* + *epidural abscess*: *primary cutaneous nocardiosis with epidural abscess, spondylodiscitis with epidural abscess* and within hyponymic relation: *phlegmon, arachnoiditis, discitis, vertebral osteomyelitis*.

These include not only meningitis but also damage to the spinal cord, nerve injury, and *an epidural abscess*, which can cause incontinence, urinary retention, fever, and back pain. (treatingbackpain.org)

By way of contrast, with the most names of diseases in question, *epidural abscess* displays particular concrete (visible, touchable) features, it is classified as countable, and the indefinite article is preferable to fulfil language users' needs; indeed, it is the context that determines the choice of the article:

Being treated for illnesses such as blood infections, sinus infections, and ear infections can help reduce your risk of developing an epidural abscess. (columbianeurology.org)

The indefinite article has been recorded as the most frequent lexical unit (48 concordances) in the first lemma to the left from the compound noun epidural abscess; in contrast to this, the definite

article is found only in 4 concordances. The morphological behaviour of the countable noun regulates the application of individual articles and the semantic references that naturally emerge from it – its descriptive or classifying roles. However, in this respect, the patterns for countable nouns are difficult to relate to the uncountable nouns patterns.

Huntington's Chorea

Out of the lexical units studied by us, the compound noun *Huntington's chorea* is the least covered in the concordances in the English Web 2015 (enTenTen15). Its etymological and stylistic peculiarities predestine it to collocate with other lexical units within the discourse only in the limited scope. In other words, there are no adjectives that would add further semantic traits to it; nouns + prepositions that occur in its closest surroundings are exclusively from the field of medicine: *a medication for, the genetic illness of, genetic testing of*. The same is valid both for the active and passive verbal structures in these positions: *investigate, inherit, treat Huntington's chorea and became hospitalized with, was treated for* (marginally, *is known as, was named, is called*).

You're right, of course. The disease was originally named '*Huntington's Chorea*' from the Greek word 'chorea' meaning 'to dance' after the involuntary jerky (dance-like) movements characteristic of the disease. (*nhsinform.scot*)

The compound noun *Huntington's chorea* is most often recorded in the surroundings of *tardive dyskinesia, Wilson's disease, Tourette syndrome, amyotrophic lateral sclerosis (ALS), multiple sclerosis, tardive dystonia*; with the help of these lexemes, semantic fields of illnesses symptoms are created. The connection between *Huntington's chorea* and other possible psychic problems of the patients is indicated by hyponymic relations with denominations like *Alzheimer's, Parkinson's, schizophrenia, depression, bipolar disorder, degeneration*.

The compound noun *Huntington's chorea* is the only one out of the ten studied lexical units that do not present any formal expressive means of the category of determination within the English Web 2015 (enTenTen15). To list relations between each entry and other entries from the research, the number of definite, indefinite, and zero articles applied with each of them has been enumerated (Table two).

Table 2. *Number of concordances containing the definite, indefinite, or zero articles with the analysed names of nervous system diseases*

Nervous system disease	Definite article	Indefinite article	Zero article
Alzheimer's disease	1 430	248	64 690
Parkinson's disease	813	243	43 053
multiple sclerosis	18	3	43 237
amyotrophic lateral sclerosis	79	13	5 230
carpal tunnel syndrome	76	23	4 018
epidural abscess	4	48	207
Huntington's chorea	0	0	169

Research interpretation

Morphological and lexico-semantic characteristics of lexical units *multiple sclerosis*, *amyotrophic lateral sclerosis*, *Parkinson's disease*, *Alzheimer's disease*, *epidural abscess*, *Huntington's chorea* and *carpal tunnel syndrome* in English Web 2015 (enTenTen15) copy general principals of grammatical system functioning of contemporary language. At the same time, they reflect the needs of its users which include various extralinguistic factors (sociological, psychological and others).

The frequency criteria lead to distinctions between the nouns, but they also point out similarities. Similarities between meanings are evident when we compare the contexts in which the implications occur, and contextual descriptions find their grounds in various factors. Individual word classes senses are distinguished through types of noun collocate, the entity being described; thus, the researcher groups the names of nervous system diseases and word classes that collocate with them. On that account, we focus attention on the statistical calculation of word classes as recorded in the position of the first lemma to the left from the analysed units (Table three).

Table 3. *Percentage representation (%) of word classes as recorded in the position the first lemma to the left from the analysed units*

Nervous system disease	Nouns	Adjectives	Verbs	Prepositions	Articles	Others
Alzheimer's disease	27.9	18.1	21.8	21.8	7.5	2.9
Parkinson's disease	30.7	19.2	16.2	23.7	8.1	2.1
multiple sclerosis	29.8	22.7	18.9	21.9	3.2	3.5
amyotrophic lateral sclerosis	31.6	8.3	33.0	19.5	5.8	1.8
carpal tunnel syndrome	24.9	34.5	9.2	21.0	9.7	0.7
epidural abscess	29.5	21.5	12.6	24.9	9.0	2.5
Huntington's chorea	45.7	25.1	9.7	16.1	0	3.4

The other related and fundamental issue made about the investigated lexical units (the first lemma to the left or the right) concerns their mutual interchangeability in numerous contexts. This phenomenon is predetermined by several factors, and it is associated with the introduced theoretical background (Moon, 2006; Halliday et al., 2004):

- most lexical units in question are semantically related;
- the users of the language do not always perceive the discrete semantic nuances of individual words;
- the real utilization of names is strongly tied to extralinguistic reality;

- its frequent occurrence confirms the centrality of this phenomenon within the investigated field.

Conclusion

After studying the lexical units to answer the research question one, it is possible to conclude that the most typical collocations in which the nervous system disease names appear are prepositional nominal phrases, adjectival phrases, and hyponymic structures.

Regarding the research question two, the analysis has shown that the adjectives before the nervous system disease names denote mostly negative subjective qualities and emotions. They are only marginally of terminological character. This fact is in concord with the presupposition that linguistic elements reflect the mental state of a language user.

The investigation within the research question three has proved that the nervous system disease names display a significant determinative power. It is noticeable that the non-specifying use of the zero articles is possible with all names of the nervous system diseases, while the definite and indefinite articles are restricted to particular manifestations of their syntactic potential and are used for a unique effect.

The methodology employed in this paper might encourage future studies in the area of corpus linguistics, focusing on the specific issues not only from general English but ESP and media discourse as well (Karpenko, 2017). One of the future research goals would be to classify our concordances according to a domain, genre, text type, or publication date; to observe whether the terminology is used differently in books addressed at a professional audience from texts written for the general public or to distinguish a possible change of meaning over the time axis.

In the present paper, the relationship between diseases of the nervous system and mental disorders has emerged into the foreground. One of the future specific research areas would be to conduct a study highlighting the mutual interconnections between the diseases mentioned above and mental health conditions like addictions, autism, bipolar disorder, depression, and others. Broader research perspectives are available when studying the linguistic phenomena in parallel (English-Slovak or English-Russian) corpora since languages influence each other semantically (Lacková et al., 2019).

The results presented above lead to several recommendations for English teachers to elaborate vocabulary in the teaching process, especially at medical universities. Furthermore, this paper might contribute to the understanding of processes that underline the choice of lexical units in given contexts, both from linguistic and psychological points of view. Therefore, the corpus analysis results have practical relevance for health practitioners and educators concerned with health issues.

About the Author:

MARTA LACKOVÁ works at the Department of English Language and Literature, Faculty of Humanities, University of Žilina. Her research focuses on comparative and corpus linguistics. She published two monographs on the grammatical category of determination and inter-lingual

homonyms, two morphological textbooks intended for university learners, and several articles.
ORCID ID is: <https://orcid.org/0000-0002-5784-7523>.

References

- Atkins, S., & Harvey, K. (2010). How to Use Corpus Linguistics in the Study of Health Communication. In T. McEnery & A. O'Keeffe. (eds.), *Corpus-based Language Studies: An Advanced Resource Book* (pp.605-619). New York: Routledge.
- Baker, P., Hardie, A., & McEnery, T. (2006). *A Glossary of Corpus Linguistics*. Edinburgh: Edinburgh University Press.
- English Web 2015 (en TenTen15). [Online] Available from: <https://app.sketchengine.eu/>.
- Fellbaum, C. (ed.). (2007). *Idioms and Collocations: Corpus-based Linguistic and Lexicographic Studies*. London and New York: Continuum.
- Flowerdew, L. (2004). The Argument for Using English Specialized Corpora to Understand Academic and Professional Language. In U. Connor & T. A. Upton (eds.), *Discourse in the Professions: Perspectives from Corpus Linguistics* (p.11-33). Amsterdam: John Benjamins.
- Friginal, E. (ed.). (2017). *Studies in Corpus-Based Sociolinguistics*. New York: Routledge.
- Gavioli, L. (2005). *Exploring Corpora for ESP Learning*. Amsterdam/Philadelphia: John Benjamins.
- Halliday, M. A. K., Čermáková, A., Teubert, W., & Yallop, C. (2004). *Lexicology and Corpus Linguistics: An Introduction*. London: Continuum.
- Jones, C., & Waller, D. (2015) *Corpus Linguistics for Grammar*. New York: Routledge.
- Kadorová, M., & Ondráčková, Z. (2005). The Use of Articles with the Names of Illnesses and Diseases in the English Language. *Insights into the Dynamics of the English Language and Culture*. 488-495.
- Kadorová, M., & Ondráčková, Z. (2005). Illnesses and Diseases: Article or Not? That's questions! *Communications*, 7, 69-78.
- Karpenko, Y. (2017) Multimedia Computer Software for Professional Training of Prospective Specialists in Foreign Languages for Preschool and Primary Education. *Information Technologies and Learning Tools*, 57 (1) 50-55.
- Lacková, M., Hundarenko, O., Moskalenko, O., & Demchenko, I. (2019). Word-formation Characteristics of Anglicisms in the Russian Slang. In *International Journal of English Linguistics*, 9(5) 283-300.
- Leláková, E. (2018). *Learning Gastronomy Vocabulary through Corpus*. Žilina: Žilinská univerzita.
- Mahlberg, M. (2005). *English General Nouns: A Corpus Theoretical Approach*. Amsterdam/Liverpool: John Benjamins.
- McEnery, T., & Hardie, A. (2012). *Corpus Linguistics: Method, Theory and Practice*. Cambridge: Cambridge University Press.
- McEnery, T., Xiao, R., & Tono, Y. (2006). *Corpus-based Language Studies: An Advanced Resource Book*. New York: Routledge.
- Moon, R. (2006). What Can a Corpus Tell us about Lexis? In T. McEnery, X. Richard, & Y. Tono. (eds.) *Corpus-based Language Studies: An Advanced Resource Book*. New York: Routledge. (pp. 197-211).
- Semeniuk O., Leleka, T., Moskalenko, O. (2020). Globalization of the Ukrainian Language Lexical System: Age, Gender and Educational Dimensions. *Advanced Education*, 15, 81-88.

- Skelton, J. R., Wearn, A. M., & Hobbs, R. F. D. (2002). A Concordance-based Study of Metaphoric Expressions used by GPs and Patients in Consultation. *British Journal of General Practice* 52(545) 114-118.
- Oxford Dictionary of English*. 2006. (3rd edn). Oxford: Oxford University Press.
- Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Schmitt, N. (2010). *Researching Vocabulary. A Vocabulary Research Manual*. New York: Palgrave Macmillan.
- Stubbs, M. (2001). *Words and Phrases: Corpus Studies of Lexical Semantics*. Oxford: Blackwell.
- Svenja, A., Brown, B., Carter, R., Crawford, P., & Sahota, O. (2004). Applied Clinical Linguistics: Corpus Linguistics in Health Care Settings. *Journal of Applied Linguistics*, 1(1), 19-28.
- Wallis, S. (2020). *Statistics in Corpus Linguistics Research*. New York: Routledge.