Tamara M. Jevrić[*]

University of Priština, Kosovska Mitrovica

Faculty of Philosophy

A CORPUS STUDY OF DERIVATIONAL MORPHOLOGY – PREFIXES UN- AND NON- IN THE BNC

Abstract: Corpora as a tool for studying morphology has been mainly used to examine morphological productivity, since English is rich in derivational morphology. Corpora can also be used to study the relationship between collocations and affixes which constitute them. The aim of this research is to establish the similarities and differences between nouns which follow adjectives with prefixes *un-* and *non-* in collocations with *unmarried* and *non-married* in the *British National Corpus* (BNC). The emphasis is on the occurrence of nouns which denote human beings. The aim is to learn what characterises the prefixes and their distribution. By focusing on the prefixes in *unmarried* and *non-married*, we also examine how an electronic corpus can help bring semantic and morphological analysis closer together, and whether it can yield significant findings about culture and society.

Key words: corpus linguistics, derivational morphology, prefixes *un-* and *non-*, BNC.

[*] tamarajevric@yahoo.com

INTRODUCTION

The use of corpora as a tool for linguistic study has proven to be productive across a variety of language fields, from lexical analysis and key concepts in semantics, to revelations about society and culture stemming from such studies. Much of the work that uses corpora as a tool to study morphology, particularly derivational morphology, seems to revolve around the issue of morphological productivity (Bybee, 2010; Baayen, 2009; Säily, 2011).

Morphological productivity refers to "the possibility to coin new complex words according to the word formation rules of a given language", cautioning us that "there is still no consensus about the nature of productivity" (Plag, 1999, p. 6). By observing the process of derivation we can notice that the productivity of affixes may vary, from those that are highly productive, to those that exhibit no productivity.[1] Advancements in understanding morphological productivity may be attributed to the use of corpora because they provide quantitative and falsifiable empirical research paradigm, without which these advancements would have been impossible (Baayen, 2009, p. 39).

Gries (2014, p. 291) sees morphology in corpus-based studies as an area whose research contributes to cognitive linguistics. This is not surprising, considering that Gries argues that corpus linguistics should be placed within the scopes of psycholinguistics, as well as cognitive linguistics, as there are many concepts in corpus linguistics which can be studied from the perspective of the first two mentioned linguistics (Jevrić, 2017b, pp. 12–13). He particularly emphasises his work on blends which "have to strike a balance between different and often conflicting facets of phonological similarity and semantics while at the same time preserving the recognizability of the two source words entering into the blend" (Gries, 2014, p. 292).

Lindquist (2009) dedicates chapters of his book to areas of research studied by means of corpora. One such chapter is that of grammar. Since grammar comprises morphology and syntax, Lindquist covers both areas. He supplies examples of how syntactic structures can be observed in corpora (constructions of passive with get, adjective complementation, etc.), but also gives one example of the employment of an electronic corpus to study morphology, that of a possible disappearance of the form *whom* (Lindquist, 2009, pp. 131–134). Inspired by Sapir's prediction that "within a couple of hundred years" "'whom' will be as delightfully archaic as the Elizabethan

---

[1] See *Rare, obscure and marginal affixes in English* (Bauer, 2014) for details on infrequent affixes.

'his' for 'its'", Lindquist uses the Time Corpus, which contains American English (AmE), to analyse the claim. By contrasting it with *who*, he shows that from the 1950s the frequency of *whom* in AmE has remained fairly stable. This he contributes to either content or writing style in which there is a stronger focus on people, i.e. a stronger use of direct questions or relative clauses pertaining to people.

Revelations about morphology can arise from studies that deal with other areas. In the paper *'Medical men' and 'Mad women' – A Study into the Frequency of Words through Collocations* (Jevrić, 2017a) the BNC is used to study the most frequent adjectival collocates before lemmas WOMAN and MAN. The adjectives are grouped into semantics fields according to their meaning. One of the groups describes marital status – *love relations and marriage*. For the lexeme *women* three adjectives are listed in the section of collocates with positive meaning – *married*, *non-married*, and *remarried*.

What makes the collocate *non-married* distinct from other collocates is that major publishers (Cambridge, Oxford, Collins Cobuild, Macmillan) do not register it as a word. *Married* changes its meaning by undergoing the process of affixation only with prefixes *un-* or *re-*. While *re-* is used mainly with verbs to indicate a repeated action, prefixes *un-* and *non-* are used to make words negative (other negative prefixes include *a-*, *de-*, *dis-*, *il-*, *im-*, *in-*, *ir-* and *no-*). Carter and McCarthy (2006, pp. 475–476, 737) list main prefixes in English explaining their meaning and exemplifying them. *Non-* is defined as not, with examples following mostly nouns after the prefix: *non-conformist*, *non-smoker*, *non-stick*, *non-believer*. *Un-* has two meanings: remove – *undress*, *undo* or reverse, not – *unhappy*, *unimportant*, *unlucky*. If we look at the same meaning of the prefixes *un-* and *non-*, that of not, the obvious pattern here is that *un-* is usually followed by adjectives, while *non-* occurs with nouns. Searching through the online *Cambridge Dictionary* the prefix *un-* is defined as being used before adjectives, adverbs, verbs, and nouns ("Un-", n.d.). *Non-* is used to make adjectives and nouns negative ("Non-", n.d.a).

Researching the etymology of the prefixes led us to the *Century Dictionary* which points to one major difference between the prefixes: *Non-* "differs from *un-* in that it denotes more negation or absence of the thing or quality, while *un-* often denotes the opposite of the thing or quality" ("Non-", n.d.b). *Un-* vs *non-* is also a matter of Old English vs Middle English, namely *un-* is an Old English prefix, while *non-* was borrowed from Romance or Latin languages during the period of Middle English causing competition between native and non-native affixes (Kastovsky, 2006, p. 169). Another important distinction is pointed out by Plag (2003, p. 100): "Negation with *non-* does not carry evaluative force" while the adjectival *un-* does.

The meaning of *married* is that it mostly describes nouns which denote human beings, men or women. This, of course, does not exclude the possibility of other collocates occurring with *married*, such as *married person*, *married people*, *married parents*, and so on. What stands in common for all these nouns, when their meaning is stripped down to basic features using means of componential analysis is that they are either male or female. Thus, nouns which follow these two prefixed adjectives can be diverse and the difference between them can reveal the attributes of affixes and their distribution. They can also explain how electronic corpora can help us understand meaning through research which combines elements of semantics, corpus linguistics and morphology.

METHODOLOGY

The subject of this research is the contrastive analysis of collocations with adjectives containing prefixes *un-* and *non-* which form the adjectives *unmarried* and *non-married* and the nouns which follow them. In English, adjectives are most commonly used in attributive position, and less likely in their appositive position (Jevrić, 2017b, p. 199). The aim of the research is to establish the similarities and differences the nouns have concerning their meaning, especially relating to the occurrence of nouns which denote human beings, in order to learn what characterises particular affixes in English and their distribution.

The corpus includes collocations with nouns which follow immediately after adjectives, when the adjective and noun are interrupted by another adjective (e.g. *unmarried Protestant woman*), or a conjunction (e.g. *unmarried and married couples*). This is enabled by the principle of collocational *span* (Lindquist, 2009, pp. 73–87), according to which collocational analysis can extend the research to five words before or after the node word. The principle stems from Sinclair's definition of collocations: "Collocation is the occurrence of two or more words within a short space of each other in a text" (Sinclair, 1991, p. 170). The corpus encompasses both adjectival and window nominal collocations, i.e. nouns that stand immediately next to the adjectival collocates, and those four to five places left of the noun.

The corpus is extracted from the *British National Corpus* (BNC) via an interface.[2] The adjectives are searched through the option *list* and then grouped based on their meaning. Since the corpus uses the program CLAWS (Constituent Likelihood Automatic Word-tagging System) to mark the part of speech of words in the corpus, we leave room for the program to incorrectly tag the adjectives. Lindquist (2009, p. 47) claims that out of 33 words in the corpus, one will be tagged incorrectly, amounting it to 97–98% of accuracy. Those words will be left out of the analysis.

ANALYSIS AND DISCUSSION

The search of *unmarried* resulted in 579 instances. Not all of them appear in collocations. The word *unmarried* is found mainly in texts belonging to the fields of social science, fictitious prose and in biographies. The nouns which follow *unmarried* are divided into four groups.

The first group includes nouns which denote human beings. Some of them are considered to be hyperonymous to other nouns in the group, or to nouns within the same semantic field the nouns are divided into. Many nouns are immediate and intermediate family members. The nouns are tabulated below:

Table 1: Nouns which collocate with *unmarried*

Semantic field	Node word (number of occurrences)
people	*males* (1), *females* (2); *person* (1), *persons* (2), *people* (3), *(young) people* (1)
partner	*counterpart* (1), *partner* (1) *partners* (2)
men and women	*man* (4), *men* (3), *(cohabiting) men* (1), *(Palestinian) men* (1), *(young) men* (1); *woman* (9), *Woman* (1), *(Edwardian middle-class young) woman* (1), *(Protestant) woman* (1), *women* (43), *women (and widows)* (1), *(young) women* (2), *(bourgeois) women* (1), *(or widowed) women* (1), *(working) women* (1)

---

[2] Davies, M. (2004-). *British National Corpus* (from Oxford University Press). Available online at: https://www.english-corpora.org/bnc/.

parents	*parent* (1), *parents* (6); *father* (6); *mother* (20), *Mother* (1), (working) mother (1), *mothers* (45), *Mothers* (2), *(teenage) mothers* (1), *mum* (2), *mums* (3), *(and uncommunicative) mum* (1)
boys and girls	*girl* (9), *(Irish) girl* (1), *girl (and a mother)* (1), *girls* (6), *Girls* (1), *(Maltese) girls* (1), *(teenage) girls* (2)
children	*child* (2), *children* (5), *(minor) children* (1); *son* (4), *sons* (1), *son or daughter* (1); *daughter* (9), *Daughter* (1), *daughters* (4), *(Clifford) daughters* (1), *(adult) daughters* (1), *(minor) daughters* (1)
grandchildren	*grandchildren* (1), *granddaughter* (1)
siblings	*siblings* (1), *brother* (2), *brothers* (2), *brothers and sisters* (1), *sister* (5), *sisters* (5)
extended family members	*nephews* (1), *(sisters and) aunts and nieces* (1), *aunt* (2)
ladies and gentlemen	*lady* (2), *(middle-aged) lady*, *(young) lady* (1), *(Scottish) lady*, *ladies* (1)
other	*Brown* (1), *concubinist* (1), *friends* (1), *heirs and heiresses* (1), *Lord* (1), *minor* (3), *owner* (1), *rape victim* (1), *(Edinburgh-)resident* (1), *teenagers* (1), *Widows and Widowers* (1)

In the first two semantic fields nouns which refer to both sexes are dominant – lemmas PERSON with one occurrence of *person*, two occurrences of *persons* and four occurrences of *people*, and PARTNER with one occurrence of *partner* and two occurrences of *partners*. In the field *men and women*, the number of lemmas which denote women outweighs the number of men, namely, the corpus generated twelve instances of *unmarried woman* and 49 instances of *unmarried women*. *Man* and *men* occur with *unmarried* four and six times in that order.

PARENT as a hyperonym has seven appearances, as the lexeme *parent* occurring one time and *parents* occurring six times. *Fathers* appears in the corpus only six times, while *mother* 22 and *mothers* 48 times. The lemma MUM is generated in the corpus search three times as *mum* and three times as *mums*. The lemma DAD did not come up in the search.

The lemma GIRL appears in the corpus eleven times as *unmarried girl* and ten times as *unmarried girls*. The lemma BOY does not appear to collocate with *unmarried*. In the field which groups sons and daughters, *daughter* and *daughters* occur eleven and seven times, while *son* and *sons* occur five and one times respectively. CHILD occurs two times as *child* and six times as *children*.

In the fields *siblings*, *grandchildren* and *extended family members*, *sisters* has seven occurrences (as an adjacent collocation, as well as in *unmarried brothers and sisters*, and *unmarried sisters and aunts and nieces*), *brothers* has three, and similarly, *sister* has five occurrences compared to two which *brother* has. There is only one occurrence of *granddaughter* in the field *grandchildren*, with no male counterpart to match. The male counterpart is also missing for *aunt*, which has three occurrences, while *unmarried nephews* and *unmarried nieces* are one each. Similarly, *lady* (both in adjacent and window collocations) has four occurrences. *Lord* in *other* refers to God. *Ladies* does not have a male counterpart.

The last field in Table 1 is *other*, and it brings together nouns which cannot be placed in other fields in the table. *Unmarried minor* is the most common collocation with three occurrences. Three nouns have a distinct meaning of male and female – *concubinist*, *Widows* and *Widowers*. *Brown* refers to a man. Two nouns are morphologically marked for gender – *heirs* and *heiresses*. There seems to be an uneven number of nouns in favour of men. However, when we take into consideration that *widows* also appears in *women and widows* and as an adjective *(widowed) women*, the number of nouns becomes more balanced, and also accentuates the correlation between marriage and widows, rather than widowers.

The second group includes nouns which denote profession. Each noun has only one occurrence: *building workers* (1), *(Finnish) farmers* (1), *office manager* (1), *(young) officers* (1), *priest* (1), *royals* (1), *schoolteacher* (1), *(female) servants* (1), *(male outdoor) staff* (1), *store assistants* (1). Two nouns are marked for gender, with the addition of an adjective – *(female) servants*, and *(male outdoor) staff*.

The third group comprises nouns which denote groups of people: *cohabitation* (1), *couple* (4), *couples* (10), *(cohabiting) couples* (1), *(and married) couples* (1), *group* (1), *(sibling) households* (1), *households* (1), *laity*[3] (1), *ones* (1), and *subjects* (1). The

---

[3] Refers to "all the people who are involved with a Church but who are not priests" ("Laity", n.d.).

most common noun within this group is *couples* which appears in texts that are thematically varied, from social sciences and religion, to newspaper reports.

The final group gathers nouns with a common meaning of *state*: *motherhood* (3), *state* (3), *condition* (1), *love* (1) and *sex* (1). The only result which clearly points to nouns specifically used in relation to women is *motherhood*. Although it has only three occurrences, *motherhood* stands out as a noun since it is found in three different sources, rather than appearing in a collocation used by one author.

What is consistent about nouns in the second group is that all nouns which denote women are larger in number. In some fields male counterparts do not even appear in the corpus results. This is a very clear indicator about different social norms pertaining to men and women. The state of being married seems to be relevant in a woman's life, regardless of age or nationality. Here, an electronic corpus can also be used as a tool that can tell us about the roles of women in society and thus provide basis for criticisms of such roles.[4]

If we were to compare two large corpora, the BNC and the internet which is, in essence, a type of electronic corpora, but also a different storage of data not compiled for linguistic analysis, a *Google* search of *unmarried* results in around 30 million examples of the word. Both corpora seem to strike a balance between the number of usages, and attest to *unmarried* as an adjectival modifier commonly used to describe nouns which denote human beings. The occurrence of *un-*, again, becomes a matter of morphological productivity. As a class-maintaining derivational affix, *un-* is highly productive (Plag, 1999, p. 113), which is exemplified by this research.

The research of the corpus yielded 19 results of *non-married*. All results of *non-married* but two appear before nouns. The nouns which follow it and the number of their occurrences are: *adults* (1), *(couple) households* (2), *man and woman* (1), *men* (1), *people* (2), *proportion* (1), *women* (9).

The most common collocation with *non-married* is *non-married women* producing nine results. Other node words are few, occurring either one time or two times. Eight of them are nouns which denote human beings, two appear with nouns in relation to humans, *non-married couple households*, and one is used in a language of calculating the number of people, "a larger *non-married* proportion" (Davies, 2004-,

---

[4] Analysis of literature also provides a foundation for such criticism, wherein woman characters can be accused of rejecting their femininity, thus exposing the expectations of men towards women (Ćuk, 2012, p. 131).

FP4[5]). In one concordance *non-married* occurs as a noun, "the portion of married to *non-married* decreases" (Davies, 2004-, CKP) and in another as an adjective with no noun to follow, "sometimes married and sometimes *non-married* separately" (Davies, 2004-, K8Y).

Since *non-married* is not registered in dictionaries as a headword, a closer examination of broader context is needed. The option *chart* reveals that 15 appearances in texts are classified as academic, one as non-academic and three as miscellaneous. Academic texts are *social sciences* texts. Sixteen occurrences of *non-married* (in collocations or otherwise) appear in a small number of books, namely, in *Women and Poverty in Britain* (10), *The Changing Population in Britain* (3), *Growing Old in the Twentieth Century* (2), *Contemporary British Society* (1), while two go under *humanities arts* category, in a book *Feminist Perspectives in Philosophy*, as collocations *non-married man* and *non-married woman* in *non-married man and woman*.

The non-academic text with the collocation *non-married women* is about social work: *Family work with elderly people*. In the *miscellaneous* category we find that one adjective appears in a book about statistics: *Interpreting Data. A First Course in Statistics*. Two collocations, *non-married adult* and *non-married women* which are also in the *miscellaneous* category, are found in texts classified as *institute doc*, *Official leaflets* which, again, deal with statistics.

Insight into the broader context as well as the actual source of collocations allows us to consider Lindquist's caveats about corpora containin "all kinds of mistakes, speech errors etc." (Lindquist, 2009, p. 10) as well as possible triviality of the findings. The non-existence of *non-married* in dictionaries would certainly compel us to do so. However, the presence of *non-married* seems to be limited to social science texts which include a high level of statistical data, where one would not expect to find examples of incorrect usages of language. Lindquist's mentioning of speech error leads us to presume that mistakes would more likely occur in speech rather than writing,[6] although separate research would be necessary to prove that claim.

Corpora do register real, authentic usages of language. If we compare the results from the BNC to *Google* search of *non-married* we notice that one and a half million results of *non-married* are generated. Not all those examples appear in collocations, but they are a testament to how some words can appear to be rare in one corpus, but

---

[5] BNC document identification codes, as Pearce (2008, p. 8) describes them.

[6] Around ninety per cent of the BNC are written texts, the rest is transcribed speech.

common in another. This discrepancy can serve as a caveat as it may point to the limits of corpus usage for linguistic research. It also shows that in terms of morphological productivity, *non-* is noticeably less productive than *un-*.

These findings bring out the issue of representativity both of electronic corpora and dictionaries. It is a common consensus that corpora can only strive to represent language by continually compiling more data, and that they can never succeed in covering all words that exist in real, authentic language usage and all possible combinations of language elements. Likewise, dictionaries are examples of the same peculiarity, but with an additional element of an inclusion of words into a dictionary being a somewhat arbitrary matter (Plag, 1999, p. 27). Covering the period between the 1980s and 1993, the BNC is certainly not representative of what language may look like now, but it might bear witness to the usage of *non-married* increasing over time and thus providing the aforementioned Google results. Another plausible explanation of the disparity of *non-married* is that the web is a significantly larger corpus, thus providing a larger degree of representativity compared to the BNC.

CONCLUSION

Since English is rich in derivational morphology this corpus study focuses on derivational affixes *un-* and *non-* in the BNC. To understand the affixes and their distribution, as well as the nouns which follow the prefixed adjectives, we examined them through the use of collocations with adjectives *unmarried* and *non-married*. Both adjectives are to a greater number, followed by nouns which denote human beings of the female sex. If we "know the word by the company it keeps" (Firth, 1957, p. 11), we can safely confirm that all women, irrespective of their differences like age, religion, etc., frequently occur in collocations that define their marital status. The BNC contains collocations which expose social norms to which women are expected to conform. The suggestion is that marriage is regarded as an important cultural construct built around ideas about women and their place in society.

A contrastive analysis of the sources in which the two prefixed adjectives appear demonstrates a common field – social sciences. The difference is how *unmarried* and *non-married* are used. *Non-married*, albeit rare, is used to represent statistical data by strictly grouping people into *married* and *non-married*. *Unmarried* is used as simply pointing to a person's marital state: "The Church of England is also responsible for some homes, but *unmarried mothers* have never been a very popular cause for charity funding" (Davies, 2004-, FU1). It is also found in literary genres of prose, "I don't know if prudent or reckless love is the better, monied or penniless love the

surer, heterosexual or homosexual love the sexier, married or *unmarried love* the stronger" (Davies, 2004-, G1X) and biographies, "It was indeed a performance to get your hair cut there as the two elderly *unmarried brothers* quite unwittingly put on a music hall act" (Davies, 2004-, B22).

A stark difference between *unmarried* and *non-married* arises when we compare the number of occurrences. *Unmarried* is significantly more common in the corpus, making the prefix *un-* a highly productive affix in the BNC, contrary to *non-*. If the prefix *non-* is defined as the absence of the thing or quality, a possible correlation with its meaning and its use in statistics is established. In the language of statistics, words with neutral prosody are employed, and *non-* has it. Further ways of research on derivational morphology could include the examination of other prefixed adjectives and potentially tying them to semantic prosody.

REFERENCES

Baayen, R. H. (2009). Corpus linguistics in morphology. Morphological productivity. In A. Lüdeling & M. Kytö (Eds.), *Corpus Linguistics* (899–919). Berlin, New York: De Gruyter.

Bauer, L. (2014). Rare, obscure and marginal affixes in English. *Lexis 8*. Retrieved from http://journals.openedition.org/lexis/304.

Bybee, J. L. (2010). *Language, Usage, and Cognition*. Cambridge: Cambridge University Press.

Carter, R. & McCarthy, M. (2006). *Cambridge grammar of English. A comprehensive guide. Spoken and written English. Grammar and usage*. Cambridge: Cambridge University Press.

Ćuk, M. (2012). Jestiva žena: „jestivi" i „kanibali" u „primitivnoj" i „civilizovanoj" mitologiji Margaret Etvud. *Reči, 4* (5), 121–136.

Davies, M. (2004-). *British National Corpus* (from Oxford University Press). Available online at: https://www.english-corpora.org/bnc/.

Firth, J. R. (1957). A synopsis of linguistic theory, 1930–1955. *Studies in Linguistic Analysis. Special Volume of the Philological Society*, 1–32.

Gries, T. S. (2014). Corpus and quantitative methods. In J. Littlemore & J. R. Taylor (Eds.), *The Bloomsbury Companion to Cognitive Linguistics* (279–300). London: Bloomsbury.

Jevrić, T. (2017a). "Medical men" and "mad women" – A study into the frequency of words through collocations. *[sic] – časopis za književnost, kulturu i književno prevođenje 2* (9). Retrieved from https: //www.sic-journal.org/ArticleView.aspx?aid=466

Jevrić, T. (2017b). *Kontrastivna analiza kolokacija sa imenicama koje označavaju ljudska bića u engleskom i srpskom jeziku* (Unpublished doctoral dissertation). Univerzitet u Beogradu: Filološki fakultet.

Kastovsky, D. (2006). Typological changes in derivational morphology. In A. van Kemenade & B. Los (Eds.), *The Handbook of the History of English* (pp. 151–176). Oxford: Blackwell Publishing Ltd.

Laity. (n.d.). In *Cambridge Dictionary online*. Retrieved from https://dictionary.cambridge.org/dictionary/english/laity

Lindquist, H. (2009). *Corpus linguistics and the description of English*. Edinburgh: Edinburgh University Press.

Non-. (n.d.a). In *Cambridge Dictionary*. Retrieved from https://dictionary.cambridge.org/spellcheck/english/?q=non-married

Non-. (n.d.e). In *Online Etymology Dictionary*. Retrieved from https://www.etymonline.com/search?q=non-

Pearce, M. (2008). Investigating the Collocational Behaviour of MAN and WOMAN in the BNC Using Sketch Engine. *Corpora, 3* (1), 1-29.

Plag, I. (1999). *Morphological productivity. Structural constraints in English derivation*. Berlin: Walter de Gruyter.

Plag, I. (2003). *Word-formation in English*. Cambridge: Cambridge University Press.

Säily, T. (2011). Variation in morphological productivity in the BNC: Sociolinguistic and methodological considerations. *Corpus Linguistics and Linguistic Theory, 7* (1), 119–141.

Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.

Un-. (n.d.). In *Cambridge Dictionary online*. Retrieved from https://dictionary.cambridge.org/dictionary/english/un?q=un-

TAMARA M. JEVRIĆ

KORPUSNO ISTRAŽIVANJE DERIVACIONE MORFOLOGIJE – PREFIKSI
UN- I *NON-* U *BRITANSKOM NACIONALNOM KORPUSU*

Rezime: Korpusi kao sredstvo izučavanja morfologije se mahom koriste za izučavanje produktivnosti, s obzirom na to da je u engleskom jeziku derivaciona morfologija veoma izražena. Korpusi se takođe mogu koristiti kako bi se analizirao odnos između kolokacija i afiksa koji ih sačinjavaju. Cilj ovog istraživanja jeste da se utvrde sličnosti i razlike između imenica koje idu posle prideva sa prefiksima *un-* i *non-* u kolokacijama sa *unmarried* i *non-married* u *Britanskom nacionalnom korpusu* (BNK). Akcenat je na imenicama koje označavaju ljudska bića. Cilj je, takođe, da se utvrdi šta karakteriše te prefikse i njihovu distribuciju. Fokusiranjem na prefikse u *unmarried* i *non-married*, izučavamo i kako elektronski korpusi mogu da približe semantičku i morfološku analizu, i da li mogu da daju značajne nalaze o kulturi i društvu.

Ključne reči: korpusna ligvistika, derivaciona morfologija, prefiksi *un-* i *non-*, *Britanski nacionalni korpus*.