

# Discourse Prosody: Semantics of Collocates of the Word *Bachelor* and Its Lithuanian Counterpart Word *Viengungis*

Linus Selmistraitis

Vytautas Magnus University, Education Academy, T. Ševčenkos St. 31, LT-03111 Vilnius, linas.selmistraitis@vdu.lt

---

**Abstract.** The current study is an attempt to identify similarities and differences of collocates with the headwords *bachelor* and *viengungis* in free word combinations and collocations of the English and Lithuanian languages and identify what the speaker's attitudes they express concerning the referent. English and Lithuanian media corpora containing free word combinations and collocations with the headword *bachelor* / *viengungis* manifest more similarities than differences in terms of semantics and connotation of collocates. However, the number of negative collocates with the Lithuanian word *viengungis* is proportionally bigger than in English which shows a more negative attitude of Lithuanian society towards unmarried men, which is expressed through the media.

---

**Keywords:** *collocate, collocations, bachelor, semantics, discourse prosody, Lithuanian, English*

---

## Introductory remarks

Paradigmatic and syntagmatic levels of a language intertwine and mutually influence one another. Lexical units from paradigms are combined into comprehensible sentences following the rules of semantic valency to build free word combinations and collocations. The word 'collocation' is derived from Latin *collocatio* [con + locare 'put together / side by side'] and it is understood as a group of words that have mutual expectancy (Jackson, 2007, 106). It means that words possess the ability to predict the likelihood of the occurrence of another word. In a free word combination, words have

no restriction in their combinability and possess greater semantic independence. Free word combinations and collocations create contexts in which lexical units realize their lexical potential, on the one hand. On the other hand, free word combinations and collocations form a semantic load of every headword, which is determined by collocates that a headword combines with. At the same time, free word combinations and collocations reflect the society's attitudes towards particular social or other phenomena, and in speech it is realised through the connotative meaning of the words.

The current study is an attempt to identify similarities and differences of meaning of collocates with a particular headword in free word combinations and collocations of two different languages and to identify what attitudes of the speaker they express concerning the headword.

The analysis zeroes in on the headword *bachelor* in English and its Lithuanian counterpart word *viengungis* with their nearest linguistic environment in which these words function as nouns<sup>1</sup>.

The research hypothesis is as follows: English and Lithuanian media corpora manifest differences in the connotative meaning of collocates in attributive collocations with the headword *bachelor* / *viengungis* in English and Lithuanian, which is predetermined by discourse prosody, and reflect the societies' attitude towards the bachelor.

The aim of the research is to characterize semantics of the collocates with *bachelor* / *viengungis* and compare the semantic load of these collocates in attributive free word combinations and collocations in English and Lithuanian taking into consideration discourse prosody. The following research objectives are set:

1) To collect free attributive word combinations and collocations with the word *bachelor* / *viengungis*, which describe the bachelor in English and Lithuanian media corpora and to build a collocational range of the word *bachelor* / *viengungis* in English and Lithuanian;

2) To identify connotations the word *bachelor* / *viengungis* acquires in combination with variety of collocates and group them into positive and negative ones;

3) To identify differences and similarities in the connotational meaning of the word *bachelor* / *viengungis* in English and Lithuanian characterising bachelors in the English and Lithuanian media.

Attributive collocations and free word combinations are combinations of words consisting of a noun which combines with an adjective. Collocations are binary, i.e. they are made up of two parts, namely, the node and the collocate. The node is hierarchically higher than the collocate because the choice of the collocate depends on the node. The node *bachelor* / *viengungis* is used in its direct meaning. The collocate is used in both

---

<sup>1</sup> Further in the text we will use two words *bachelor* / *viengungis* to denote both the English headword *bachelor* and its Lithuanian counterpart word *viengungis* which will be considered in the text as one unit and used in the singular form.

direct and figurative meanings. The collocate is unpredictable and differs in English and Lithuanian.

The degree of how widely the word or the headword can combine with another word can vary greatly. That is why it is difficult to draw a clear line between free word combinations and collocations.

The collocational range embraces all the words that the word *bachelor* / *viengungis* combines with in the corpus under analysis. From three types of lexical combinations – free word combinations, collocations and idioms – the first two types are used in the research. In terms of data type, the research is based on one of the types of a corpus from the three possible (the Web as a corpus, a balanced corpus and a monitor corpus), i.e. the Web as a corpus.

For the research, the Web as a corpus will be exploited because it has the value as linguistic resource. The latter idea is supported by Kilgariff's challenging and controversial statement made at the beginning of the 21<sup>st</sup> century: "The corpus of the new millennium is the Web" (Kilgariff, 2001, 345). From the current perspective, this statement seems to be not as controversial as it used to be.

In the research, the written mode of language is used, which presents the smallest technical challenge to construct as compared to spoken or video corpora. The texts have not been annotated, i.e. the linguistic analysis has not been included in the data itself. As McEnery and Hardy put it, "Those corpora which have already been analysed in some way are annotated, those which have yet to be analysed are not" (2012, 14). The Web as a corpus does not have any encoding or corpus annotation. Thus, the principle of annotation doing it manually and selecting attributive collocations from the Web as corpus is applied in the research.

In our research, we focus on a synchronic comparison of English and Lithuanian corpora. Such factors as comparability, balance and representativeness are important in any researcher using corpora. However, as Durrell puts it, "It is best to recognize that these goals are not an all-or-nothing: there is a scale of representativity, of balancedness, of comparability. We should seek to define realistically attainable positions on these scales, rather than abandon them altogether" (as cited in Gipper & Gehrke, 2015, 17).

## Literature Review

### **Corpus linguistics and its application in studies of free word combinations and collocations**

The results of the application of findings based on corpus linguistics throw a new light on long-standing and novel branches of linguistics. Corpus linguistics is not a traditional branch of linguistics, which at first glance might be presumed as such by a nonprofessional who comes across this term for the first time. This presumption

may arise from analogical terms such as, say, cognitive linguistics or areal linguistics. Corpus linguistics does not deal with a study of a particular aspect of language. As McEnery and Hardie put it, corpus linguistics is rather “an area which focuses upon a set of procedures, or methods, for studying language” (2012, 1). However, some linguists advocate that corpus linguistics cannot be treated as a group of methods or procedures, and they definitely consider it a discrete branch of linguistics. Considering corpus linguistics as a set of methods for studying language, in the current study we used a corpus to extract a language model.

From three possible variants of data type, monitor corpora continually expand and include ever-growing text size in the course of time. As McEnery and Hardie put it, monitor corpora “could be said to balance any need to be precise about the composition of a corpus against sheer size – as the corpus grows, we might assume that any skew in the data naturally self-corrects, since there is no consistent skew in the data input” (McEnery & Hardy, 2012, 6–7). A balanced corpus represents a particular type of language over a specific span of time and this corpus is balanced and representative within a particular sampling frame which defines the type of language and population that could be characterized. The Web as a corpus is in a way similar to the monitor corpus because it is ever-growing and the data are applicable for research with interfaces that are specifically designed for the use of the Web.

The Web gives access to huge quantities of text that “presents the most provocative questions about the nature of language” (Kilgarriff, 2001, 344). We admit that the corpus from the Web contains errors and cannot be compared in this respect to processed data. However, the most recent data from the Web may give new evidence to the questions under research, having in mind that even such standard search engines as Google can help to explore the Web as a corpus. As McEnery and Hardie put it, “there is little doubt that the thousands of examples from Google would allow a more nuanced investigation of” (McEnery & Hardie, 2012, 8) a particular issue than only few examples in the BNC. So the Web as a corpus is “invaluable in cases where you need a large quantity of data in order to deal with a low frequency of occurrence” (ibid).

From the descriptive point of view, the Web as a corpus presents authentic material and reflects a natural way of language usage predominantly by the native speakers. However, we have to have in mind that “authentic does not necessarily mean reliable. Indeed, authenticity is often related to problems of ‘authoritativeness’ and it is everyday experience that authentic in the Web often means inaccurate (misspelt words, grammar mistakes, improper usage by non-native speakers), owing to its nature as an unsupervised unedited collection of texts” (Gatto, 2009, 38–39).

Since 2005, scientific literature on the Web as a corpus started to gain its power because this year was marked by the annual Corpus Linguistics Conference, which had a separate workshop on the Web as a tool for corpus linguistic research. As Lampert puts it, “Corpus linguists are at present found to revise their reservations against the

Web and in fact increasingly turn to it as their prime source of language data (Lampert, 2009, 155).

On the one hand, the Web as a corpus is the largest corpus; on the other hand, it is the messiest and anarchic one. Although the Web search engines are excellently efficient pieces of technology giving thousands of samples at one click, they present linguistic data in a rather distorted way, for example, the searched terms in headlines and titles go to the top of the list. The statistics on the frequency of the usage of a particular linguistic term is not reliable inasmuch it varies from search engine to search engine.

Kilgarriff and Greffentette's (2003) ideas about the Web as a corpus have to be supported notwithstanding all the drawbacks that texts on the Web have. The authors suggest avoiding application of value criterion while identifying what corpus is and what is not and do not mix the question "What is a corpus?" with the question "What is a good corpus?" as McEnery and Wilson do when they claim that a good corpus has to be representative, finite and machine-readable (McEnery and Wilson, 2001, 21). Kilgarriff and Grefensette (2003) set aside the semantic value of the question and suggest a broad definition of a corpus. They define the corpus as "simply as a collection of texts. If that seems too broad, the one qualification we allow relates to the domains and context in which the word is used rather than its denotation: a corpus is a collection of texts when considered as an object of language or literary study" (Kilgarriff & Grefensette, 2003, 2).

We have to admit that the Web as a corpus has a drawback because it is not permanent and finite: it cannot be downloaded and archived as ready-made texts. However, the Web as a corpus has a pragmatic advantage over corpora as a body: it is free or low cost, has open access, and is big in size. The dichotomy of corpora is evident – a corpus as a body and a corpus as the Web. The corpus as a body is characterized by such parameters as permanence, stability of data, finite size, and balance, while the corpus as the Web tends to be flexible, non-finite and non-predictable.

## **Semantic preference and semantic prosody**

In the past two decades of the 21st century, semantic prosody and semantic preference have become the two linguistic terms, which are in focus of semantists who apply corpus data for the analysis of the word semantics. As Begagić puts it, "There are almost no works on semantic preference and semantic prosody outside the scope of corpus linguistics, as empirical data used from corpora enable the linguist to make statements that are objective and based on natural language" (Begagić, 2013, 403).

Sinclair (1991, as cited in Hu, 2015), Louw (1993), Stubbs (2001), and Partington (2004) developed a theory of semantic prosody. Sinclair's theoretical underpinnings of semantic prosody are related to pragmatics. Sinclair considered "semantic prosody

to be the discourse function of a sequence rather than the property of a word” (as cited in Hu, 2015, 118). Stubbs (2001) further developed the idea and claimed that there are semantic relations between the node (headword) and collocates, and among collocates themselves, reformulated the concept of semantic prosody and called it as ‘discourse prosody’. Stubbs considered the nearest corpus environment of particular words as discourse prosody and classified discourse prosody into negative, positive or neutral. Semantic preference which is loosely related to semantic prosody deals with semantic categories shared by the frequent collocates of a specific node item (Hunston, 2007; Partington, 2004). As Hu puts it, “despite the fact that there is a fuzzy boundary between the two concepts of semantic preference and semantic prosody, one clear-cut distinction between them is that semantic preference may be in favour of any definable semantic field, but semantic prosody is always either for positive or for negative evaluation” (McEnery & Hardie, 2012, as cited in Hu, 2015, 118). In other words, as Partington puts it, semantic prosody “evaluates the topic and indicates to the hearer how a part of the utterance is to be interpreted functionally” (Partington, 2004, 149).

## Methods and Procedure

Qualitative approach is applied to investigate the data retrieved from the corpus following the principles of corpus linguistics. Descriptive method is used to identify the semantic nature of collocates with the headword *bachelor* / *viengungis* predetermined by discourse prosody.

For that purpose, free word combinations and collocations with the headword *bachelor* / *viengungis* were selected from the Web as a corpus. Google search engine was used to extract free word combinations and collocations from mass media texts in the period from 2012 to 2017. After searching mass media corpus on Google platform, all possible collocates of *bachelor* / *viengungis* were collected. 48 different collocates of the node *bachelor* and 34 collocates of the node *viengungis* were collected. All free combinations and collocations with the headword *bachelor* / *viengungis* were juxtaposed to identify differences in the semantic load of collocates which reflect the society’s attitudes towards the bachelor. All the occurrences were manually examined and observed at the span of at least 10 words to the left and 10 words to the right, and the results were compared.

The usage of corpora is based on the principle of total accountability as it is understood by Leech (Leech, 1992). Total accountability presupposes that the researcher must not select favourable sets of data, i.e. all relevant to the research data are used approaching the hypothesis. When there are too many data using the entire corpus, Leech suggests avoiding conscious selection of data which could distort the results or using unbiased (randomized) examples of the corpora. In other words, we avoid motivated selection of examples that fit the hypothesis. Of course, the corpus is also a

relative piece of data about the language as a system, so we have to make a limitation concerning the results applicable to the entire language itself. However, we should not screen out inconvenient examples or motivate our choices on the base of convenient examples proving our hypothesis. Here we have to admit that 100 percent accountability is impossible inasmuch as corpora is also only a selection from available linguistic repertoire and concordances or collocations found on the internet or in official corpus are not finite. We also have to have in mind that in figurative speech where valency of words and their semantics are often disrupted and it gives infinite number of possible collocations with a particular word that a researcher may focus on. Thus, we only strive for total accountability because corpus is finite, and language is not.

## Discussion

English dictionaries present different semantic structure of the word *bachelor*. The number of its meanings range from two to four. Exhaustive semantic structure of the word *bachelor* is given by *The Concise Oxford Dictionary* (COD, 1999): “1) a man who is not and has never been married; 2) a male bird or mammal prevented from breeding by a dominant male; 3) a person who holds a first degree from a university or other academic institution (only in titles or set expressions; 4) a young knight serving under another’s banner” (COD, 1999, 96). The origin of the word: “Middle English; from Old French: *bachelor* “a young man aspiring to knighthood” (COD, 1999, 96). Only the first meaning of the word *bachelor* will be investigated in the research. The meaning of the Lithuanian *viengungis* is as follows: “nevedęs, pagyvenęs vyras” (not married, elderly man<sup>2</sup>) (Dabartinės lietuvių kalbos žodynas, 1993). The definition says that in its semantic structure Lithuanian *viengungis* has the seme ‘being elderly’. However, more modern dictionaries give a modified definition: “Kas gyvena vienas, neturi šeimos (ppr. nevedęs, pagyvenęs vyras)” (The one who lives alone, has no family, usually an elderly not married man) (Lietuvių kalbos žodynas, 2005). The latter meaning has a semantic shift from *viengungis* as an elderly person to *viengungis* as usually an elderly person, i.e. the seme ‘elderly’ is not necessarily present. Thus, the definitions of the English *bachelor* and the Lithuanian *viengungis* are different in terms of one seme: Lithuanian *bachelor* includes the seme ‘elderly’.

The analysis of the corpora has revealed that in English there are more collocates which go with the word *bachelor* than in Lithuanian with its counterpart word *viengungis*. Correspondingly, there are 48 English collocates and 33 Lithuanian ones. The difference in the variety of collocates might be explained by the fact that topics in

---

<sup>2</sup> Translated by the author, here and further in the text

which the bachelor is described in the English media are more exhaustive than in the Lithuanian media. Moreover, the spectrum of English mass media articles and other digital publicist texts are broader that is why more diverse collocations can be found in English. What is more, the size of the corpus in English is bigger than in Lithuanian.

With reference to Partington's (2004) theory that words have either a positive, negative or neutral connotation, all English and Lithuanian free word combinations and collocations with the headword *bachelor* / *viengungis* were classified from this vantage point. There were cases difficult to ascribe as either having positive or negative features. The crucial role in this case was assigned to the context, which defines the exact connotation of the word. Still, some subjectivity in this classification was the limitation of the research.

19 English collocates which added a positive connotation to the word *bachelor* were found in the English corpus: *active, affable, attractive, charming, chaste, contented, eligible, entrepreneurial, family-oriented, handsome, optimistic, outgoing, respectable, rich, steadfast, successful, smart, tall, and young*.

A part of these collocates give positive features of the bachelor's appearance or character, e.g.:

(1) He is ***charming, outgoing*** and a lifelong bachelor.

Some collocates indicate the quality of life or inner state of the bachelor, for example:

(2) *The image of the **successful, entrepreneurial** bachelor is one that's forever etched onto our eyeballs.*

(3) *A majority of Grant's popular films in the 1990s followed a similar plot that captured an **optimistic** bachelor experiencing a series of embarrassing incidents.*

Collocates can render either positive or neutral evaluation depending whether the free word combination or collocation is uttered by a woman or a man. For example, *family-oriented* in the following discourse produced by a woman acquires a positive connotation because it goes with other adjectives, which indicate positive features by default, e.g.:

(4) *Our sophisticated, smart and active Bachelorette is seeking a **family-oriented, successful, and handsome** bachelor.*

16 neutral collocates of the node *bachelor* were identified: *black, Caucasian, dark, confirmed, dead, elderly, intense, gay, homosexual, heterosexual, lifelong, light-skinned, life-long, middle-aged, natural, and single*.

Neutral collocates identify belonging of a person to a particular ethnic group or characterize its appearance, e.g.:

(5) *Hello, I'm a single down-to-earth, professional **Caucasian** bachelor.*

It can also give no evaluation just stating the fact that the man is not married and he is not going to, e.g.:

(6) *Clooney is engaged to human rights lawyer Amal Alamuddin has given hope to women that it is possible to tie down a **confirmed** bachelor.*

(7) *A **lifelong** bachelor, Canadian prime minister Mackenzie King was obsessed with his deceased mother.*

(8) *In many ways I think Prince Charles is a **natural bachelor** but I can't believe that the princess, who is now 31, will be happy to spend the rest of her life on her own.*

The neutral collocates also give factual information about the bachelor without connotations, e.g.:

(9) *Relatives ransack **dead bachelor's** home for goodies.*

Collocates indicate sexual orientation of the bachelor and usually the context in which this type of free collocation function is neutral. For example:

(10) *The world's most in-demand **gay** bachelor has been revealed.*

(11) *We could say that a **homosexual** bachelor is more of a bachelor than a priest.*

(12) *However, a 36-year-old **heterosexual** bachelor – whom we shall have to call Bill – who works in the media, claimed he had little experience of people assuming he was gay.*

Negative connotations were supported by 13 English words: *bored, chaotic, cynical, elusive, hapless, heartbroken, hurt, irresponsible, jaded, lascivious, lonely, madcap, and married.*

They characterize the bachelor as having negative character or features of negative behaviour, e.g.:

(13) *Two years later, Hugh Grant starred in *About a Boy*, another hit, following the redemption of a **cynical** bachelor through a friendship.*

Collocates also define the bachelor as an unsuccessful one due to life circumstances, e.g.:

(14) *What has a jaded middle-aged bachelor got to look forward to?*

There are cases when the collocation with the word *bachelor* becomes an oxymoron and contains some irony. Usually, it obtains a negative collocation. For example:

(15) **Married**-*bachelor syndrome: when men fail to grow up.*

12 collocates which added a positive connotation to the bachelor were found in the Lithuanian corpus: *garsus, geidžiamas, įsimylėjęs, karščiausias, optimistiškas, patrauklus, perspektyvus, rafinuotas, tvarkingas, turtingas, žavus, and žymus*. They give positive features of appearance, character or quality of life like in the English corpus, e.g.:

(16) *Žavingas viengungis G. Clooney šiomis dienomis atrodo neįtikimai laimingas.*

Some collocates with positive connotations characterize the bachelor in positive way due to their social status or recognition, e.g.:

(17) *Žymaus viengungio širdį lietuvė bando užkariauti skaistumu.*

7 neutral collocates of the node *bachelor* were identified: *anoniminis, jaunas, paskutinis, pusamžis, statistinis, trisdešimtmetis, and vedęs*.

Similar to English, the collocates can give no evaluation and just state the fact that a man is not married (for example, belonging of a person to a particular age group), e.g.:

(18) *Per vieno draugo vestuves pabroliai susilažino, kad **paskutinis** viengungis gaus tam tikrą sumą.*

(19) *Dažnas **pusamžis** viengungis mano, kad sulaukus penkiasdešimties jam gyvenimas tik presided.*

(20) ***Trisdešimtmečiai** viengungiai neretai turi asmenybės, bendravimo, seksologinių, sveikatos problemų, tad lieka vedusieji.*

However, some of these collocates can function in a negative context and characterize the bachelor in a negative way, e.g.:

(21) ***Trisdešimtmečiai** viengungiai neretai turi asmenybės, bendravimo, seksologinių, sveikatos problemų, tad lieka vedusieji.*

Differently from the English corpus, the Lithuanian corpus does not contain neutral collocates which describe the bachelor's ethnic background or sexual orientation.

Negative connotations were supported by 14 Lithuanian collocates: *amžinas, ekstravagantiškiausias, keistuolis, mistinis, nepataisomas, pasimetęs, prisiekęs, problemiškausias, pašėlęs, senas, senstelėjęs, tingiausias, užkietėjęs*, and *vienišas*.

The same as in English, negative connotation collocates characterize the bachelor as having negative character or behaviour features, e.g.:

(22) *Atrodytų, virtas kiaušinis – tai patiekalas, kurį gali pasigaminti net pats **tingiausias** viengungis.*

(23) *Filme „Pusryčiai pas tėtį“ – **pašėlęs** viengungis, žavi mergytė ir gyvenimas, apvirtęs aukštyn kojomis.*

Differently from the English corpus, the bachelor who is not married for a long time is characterized in a negative way, e.g.:

(24) *Pirmasis **užkietėjusio** viengungio gyvenimo principas – palaikyti atsitiktinius lytinius santykius.*

(25) *2011 m. tėvu tapo dar vienas **nepataisomas** viengungis.*

(26) ***Amžinas** viengungis sieloje, galintis lakstyti nuo žmonos prie meilužės ir atgal kol visiškai neišsikvėps.*

Some of the collocates have a negative connotation which is diminished by the positive context, e.g.:

(27) *Senstelėjęs viengungis vilki *elegantiškus* kostiumus.*

However, the Lithuanian corpus does not contain collocates which would identify the bachelor in a negative way due to life circumstances.

## Conclusions

48 English collocates and 33 Lithuanian ones of the node *bachelor / viengungis* in the collected corpora were found. The headword *bachelor / viengungis* with the collocates was used in all the three types of discourse according to discourse prosody, i.e. positive, neutral and negative: in English 19 positive, 16 neutral, 13 negative; in Lithuanian 12 positive, 7 neutral, 14 negative.

In both English and Lithuanian corpora, the bachelor with collocates having a positive denotation was characterised in terms of appearance, character, quality of life or inner state.

In both English and Lithuanian corpora neutral collocates give factual information about the bachelor without any connotations, for example, belonging of a person to a particular age group. However, only in the English corpus neutral collocates accompanied the headword *bachelor* when the context of usage was related to the bachelor's belonging to a particular ethnic group or its appearance. Only in English corpora, neutral collocates indicate sexual orientation of the bachelor and usually the context in which this type of free collocation function is also neutral.

In both English and Lithuanian corpora, negative connotation collocates characterize the bachelor as having both negative character or behaviour features. Differently from the English corpus, the bachelor who is not married for a long time is characterized in a negative way in the Lithuanian media corpus.

The English and Lithuanian media corpora containing free word combinations and collocations with the headword *bachelor* / *viengungis* manifest more similarities than differences in terms of semantics and connotation of collocates resulting into negative or positive discourse prosody. However, the number of negative collocates with the Lithuanian word *viengungis* is proportionally bigger than in English which shows a more negative attitude of Lithuanian society towards an unmarried man, which is expressed through the media.

## References

- Begagić, M. (2013). Semantic Preference and Semantic Prosody of the Collocation 'Make Sense'. *Jezikoslovlje*, 14, 2–3, 403–416. Retrieved from: <https://hrcak.srce.hr/112198>.
- COD – *The Concise Oxford Dictionary* (1999). Oxford: Clarendon Press.
- Dabartinės lietuvių kalbos žodynas* (vyr. redaktorius S. Keinys). (1993). Vilnius: Mokslo ir enciklopedijų leidykla.
- Durrell, M. (2015). 'Representativeness', 'Bad Data', and Legitimate Expectations. What Can an Electronic Historical Corpus Tell us that We Didn't Actually Know Already (and How)? In J. Gipper, R. Gehrke (Eds.). *Historical Corpora. Challenges and Perspectives* (pp. 13–34). Narrverlag.
- Gatto, M. (2009). The 'body' and the "Web": The Web as Corpus 10 Years on. *ICAME Journal*, 35, 35–58. Retrieved from: [http://clu.uni.no/icame/ij35/Maristella\\_Gatto.pdf](http://clu.uni.no/icame/ij35/Maristella_Gatto.pdf)
- Hu, M. (2015). A Semantic Prosody Analysis of Three Adjective Synonymous Pairs in COCA. *Journal of Language and Linguistic Studies*, 11, 2, 117–131. Retrieved from: <http://www.jlls.org/index.php/jlls/article/view/418>

- Hunston, S. (2007). Semantic prosody revisited. *International Journal of Corpus Linguistics*, 12 (2), 249–268. <https://doi.org/10.1075/ijcl.12.2>
- Jackson, H. (2007). *Words, Meaning, and Vocabulary*. London: Continuum.
- Kilgarriff, A. (2001). Web as Corpus. In P. Rayson, A. Wilson, T. McEnery, A. Hardie, S. Khoja (Eds.), *Proceedings of the Corpus Linguistics 2001 Conference* (pp. 342–344). UCREL, Lancaster. Retrieved from: <http://ucrel.lancs.ac.uk/publications/CL2003/CL2001%20conference/papers/kilgarri.pdf>
- Lampert, M. (2009). *Attention and Recombinance. A Cognitive-Semantic Investigation into Morphological Compositionality in English*. Peter Lang.
- Leech, G. (1992). Corpora and Theories of Linguistic Performance. In J. Startvik (Ed.), *Directions in Corpus Linguistics* (pp. 105–122). Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110867275.105>
- Louw, W. E. (1993). Irony in the Text or Insincerity in the Writer? In M. Baker (Ed.), *The Diagnostic Potential of Semantic Prosodies. Text and Technology: In Honour of John Sinclair* (pp.157–176). Amsterdam: John Benjamins.
- McEnery, T., Hardie, A. (2012). *Corpus Linguistics*. Cambridge: Cambridge University Press.
- McEnery, T., Wilson, A. (2001). *Corpus Linguistics*. Edinburgh: Edinburgh University Press.
- Kilgarriff, A., Grefenstette, G. (2003). Web as Corpus. *Computational Linguistics*, V, N, 1–15. Retrieved from: <https://www.kilgarriff.co.uk/Publications/2003-KilgGrefenstette-WACIntro.pdf>
- Lietuvių kalbos žodynas: elektroninis variantas*. (2005). (G. Naktinienė (vyr. redaktorė), J. Paulauskas, R. Petrokienė, V. Vitkauskas, J. Zabarskaitė (red.)). Vilnius: Lietuvių kalbos institutas. Retrieved from: [www.lkz.lt](http://www.lkz.lt).
- Partington, A. (2004). Utterly Content in each other's Company: Semantic Prosody and Semantic Preference. *International Journal of Corpus Linguistics*, 9 (1), 131–156. <http://dx.doi.org/10.1075/ijcl.9.1>.
- Stubbs, M. (2001). *Words and Phrases: Corpus Studies of Lexical Semantics*. Oxford: Blackwell.

---

## Diskurso prozodija: žodžio *bachelor* ir jo lietuviško atitikmens *viengungis* kolokatų semantika

**Linus Selmistraitis**

Vytauto Didžiojo universitetas, Švietimo akademija, T. Ševčenkos g. 31, LT-03111 Vilnius, [linus.selmistraitis@vdu.lt](mailto:linus.selmistraitis@vdu.lt)

---

### Santrauka

Straipsnyje pateiktas tyrimas atskleidžia anglišių ir lietuvišų kolokatų su anglišku žodžiu *bachelor* ir jo lietuvišku atitikmeniu *viengungis* semantikos panašumus ir skirtumus laisvuosiuose

žodžių junginiuose ir kolokacijose, kurios parodo visuomenės požiūrį į nevedusius vyrus. Tyrimo šaltinis – 2012–2017 m. angliškos ir lietuviškos žiniasklaidos straipsniai. Tyrimas atskleidė, kad angliška ir lietuviška žiniasklaida turi daugiau panašumų nei skirtumų žodžių *bachelor* ir *viengungis* diskurso prozodijos atžvilgiu, kai kalbama apie neigiamame arba teigiamame diskurse vartojamus kolokatus su žodžiais *bachelor* / *viengungis*. Tačiau neigiamų kolokatų su žodžiu *viengungis* lietuvių žiniasklaidoje proporcingai yra daugiau nei anglų kalboje su žodžiu *bachelor*. Tai rodo labiau neigiamą Lietuvos visuomenės požiūrį į nevedusius vyrus.

---

***Esminiai žodžiai:*** kolokatas, žodžių junginys, *viengungis*, semantika, diskurso prozodija, lietuvių kalba, anglų kalba.

---

Gauta 2018 07 19 / Received 07 19 2018  
Priimta 2018 12 03 / Accepted 03 12 2018