# THE PATTERN OF  WORD USAGE VIEWED BY CORPUS LINGUISTICS

RŪTA MARCINKEVIČIENĖ

## Introduction

The advent of corpus linguistics has challenged some of the traditional views on language and its structure. One of the best examples illustrating the impact of corpus linguistics on the theory of language is the distinction between a *word* and a *lexical item.* Until the last decade the assumption that the word is a unit of lexical meaning was taken for granted. While extracting the lexical information from multi-million word corpora, lexicographers, detected strong, recurrent patterns in the way words were used in texts. Two basic facts were detected:

> a)" many, if not most, meanings require the presence of more than one word for their normal realisation;
>> b) patterns of co-selection among words, which are much stronger than any description has yet allowed for, have a direct connection with meaning." (Sinclair 1998:3).

It has been argued by researchers that meaning is created in collocation, it arises predominantly from the textual environment. Moreover, corpus linguists pursuaded "… to put aside some well-respected assumptions about language that may be hampering our thinking - such as the imbalance at present in favour of the independence of the word, rather than its interdependence on its cotext, or verbal environment."(Sinclair 1998:6). Thus corpus linguistics called into question the way lexical and semantic studies have been dominated by single words. Since the *word* is not the core item for the description of meaning, it gave way to *lexical item* meaning a unit of description made up of words and phrases with an internal structure.

Obviously, the proposed alternative model of meaning is not universally applicable. It works better  within the framework of the *phraseological tendency* of language or the so-called the *idiom principle* of text construction rather than within the opposite, i.e. *terminological tendency* or *open-choice principle* (Sinclair 1998, 1999).  The first has to do with ordinary words that are context-dependant and change their meaning mainly through gradual movements of collocation, the second concerns terms which are protected as far as possible from the effects of usage.  Thus it has been stated that two lexicons are used by speakers of a language. One is essentially an extended term bank, containing words and phrases which have fixed meanings  and clear meaning differentiation. The other type of lexicon is usage-based. It is called *the empty lexicon*. For this lexicon lexical items are not always single words. Their meaning is dependant on both a form of the lexical item and its environment (Sinclair 1996 A).

The alternative model of language is accepted as a theoretical background here, therefore usage becomes of paramount importance. Corpus linguistics addresses a hitherto largely unnoticed layer of linguistic analysis between syntax and the lexicon and corpus analysis is the method that can help us to see the patterns for what they are.

The patterns that emerge in this layer have been called differently by different authors: *patterns of word behaviour* (Levin et al 1997), *relational patterns* (Williams 1998), *patterns of usage, co-occurrence patterns, association patterns* (Biber 1993, 1996), *context profile, semantic coagulation* (Teubert 1999), *patterns of word use, norms of usage, statistically significant recurring pattern of behaviour* (Hanks 1998), *lexical sets, conventional uses* (Hanks 1996), *word association norms* (Church and Hanks 1990), *collocational and colligational patterns* (Sinclair 1996).

The variety of names reflect the diversity of aspects and viewpoints on the subject. Nevertheless, some common preliminaries in the analysis of word usage patterns are obligatory. Corpus evidence in the form of a concordance is one of them. Thus corpus evidence has to be considered in order to capture and describe *usage patterns* (this particular term has been chosen in the present paper for the convenient abstractness of its notion).

The aim of this paper is to discuss the existing descriptions of usage patterns and to offer both automated and manual procedures for distinguishing relevant features of linguistic environment from mere "noise". Besides, methodology of data extraction and processing is offered and appropriate levels of analysis and synthesis are chosen. The procedure is exemplified by the analysis of the concordance of the noun *mind* (used in singular) which has resulted in its usage pattern. *Mind* has been chosen as one of the most frequent words of English and a typical member of the empty lexicon. The size of its concordance was really challenging and provided a good opportunity to deal withone of the common words that are usually neglected by most authors.

### Different ways to present usage patterns

Usage patterns have been either mentioned or described in two kinds of publications: research papers and lexicographic presentations. Both types deal with individual lexical items but differ in the depth of analysis and delicacy of description. Research papers usually offer accurate description of selected lexical items that indicate semantic and syntactic phenomena identified in the lexical analysis of corpus data. The common features of corpus-based analytical descriptions are as follows:

- identification and characterisation of fixed lexical expressions, co-occurrence with particular words, i.e. collocation (Altenberg 1991, Kjellmer 1991, Williams 1998 among many others),
- co-occurrence with grammatical features, syntactic patterns in particular, i.e. colligation (Hanks 1996).
- co-occurrence with semantic features of the context and its pragmatic function, i.e. semantic preference and semantic prosody and (Hanks 1996, Louw 1993). All three approaches are interconnected into a full-fledged description of the meaning of a few lexical items in Sinclair 1996 B, Sinclair 1998, 1999)

The outcome of a corpus-based lexical analysis of some lexical items, beside its description and explanation, is an overview of the usage  as well as attempts to produce a lexicographic presentation of the lexical items under consideration. A generalisation, for example, of the collocation *true feelings* is as follows:

"Our conclusion is that the collocation *true feelings* is the core of a compound lexical item which has the following inherent components: a semantic prosody of reluctance-inability, a semantic preference of expression (and a strong colligation of a verb with the semantic preference), a colligating possessive adjective and the core.

So, not only are our true feelings our genuine emotions, but we use this particular collocation when talking about our reluctance to express them, even to ourselves. The collocation is almost never used except as part of this compound lexical item." (Sinclair 1996 B:90).

A more lexicographically oriented presentation of *enthusiasm* is given in Hanks (1998:160) where norms of usage are presented in the form of natural language sentences (see 1, 7 below) or as a lexicographer's comment supplied with word combinations (as in 8). Here is a selected presentation of 3 norms out of 12:

" 1. *A person may say something (or do something)* **with enthusiasm.**

7. *Enthusiasm may be **tempered by** something else (**caution, realism, cynicism,** or **a state of affairs**).*

8. *'Enthusiasm' is often used with an intensifying adjective, e.g.* **great enthusiasm, tremendous enthusiasm, enormous enthusiasm, immense enthusiasm, etc.**"

While summing up the researchers' attempts to present usage patterns it should be said that no common approach to the solution of the problem has been found. Nevertheless it is conceived that corpora offer a good opportunity to detect and describe those patterns that were unattainable to introspective method or traditional lexicography.

Lexicographic presentations, given in modern dictionaries, differ according to the amount of analysis and synthesis on the part of their compilers. One possibility to present a usage pattern of a lexical item is to select concordance lines representing typical usages. This method of presentation is valuable for the real life examples. Full sentences are given and selection is made on the level of inclusion/noninclusion of lines without any transformation of morphological or syntactic structures. On the other hand, the overall pattern is obscured if one line per collocate has to be selected and considerable part of information is omitted. Nevertheless, huge concordances cannot be used efficiently. This way of presentation takes too much space and is suitable only to electronic dictionaries like the Collins Cobuild English Collocations on CD-ROM. It provides top 20 collocates together with their concordance lines that are given for each of 10 000 node words representing most of the common words of English.

Another way of presentation consists of a list of collocates enumerated and classified according to their significance, morphological characteristics or semantic features. This is how a more detailed picture of the collocates can be obtained, however, it lacks information on the types of syntactic relationship and therefore cannot represent a pattern of usage. This method is favoured by traditional paper dictionaries of collocations such as the BBI Combinatory Dictionary of English.

The greatest amount of effort is necessary for the type of usage pattern which presents not only collocates of the word under consideration but also syntactic patterns of their

usage in the form of phrases. This is the way how corpus-based explanatory dictionaries like Collins Cobuild English Dictionary and Longman Dictionary of Contemporary English present the nominal use of *mind*. The only shortcoming of this type of presentation is as follows: the cited phrases reflect only the immediate surroundings and the most significant collocates of the node word. The syntactic patterning as well as the collocates in a more distant syntactic positions are not given.

Our aim while compiling a pattern of usage is to bridge the gap between research vs. dictionary type of description of a lexical item on the one hand, and corpus data vs. dictionary entries, on the other. Thus it should be shorter than a concordance or detailed description but longer and more detailed than an entry in a dictionary. In any case it should give a very clear idea about a word and its context.

### Methodology of corpus analysis and pattern synthesis

The detailed analysis of the usage of the singular form of the noun *mind* is based on the concordance of 54 580 lines, derived from the Bank of English. At the time of data extraction the Bank of English contained ca 300 million words. The aim of the analysis of the concordance is threefold: to identify, single out and to present collocations in a way which would reflect the pattern of usage is.

The first step in the analysis of the corpus data is identification of collocations with the core item *mind*. Corpus linguists describe several distinctive features of a collocation based on a variety of criteria (see Marcinkevičienė 1995 and 1999 for an overview). Out of the numerous criteria for the identification of a collocation only three have been selected for practical purposes, i.e. the frequency of co-occurrence, meaningfulness and the presence of at least two notional words as constituents.

The first criterion poses some problems: what method should be used for finding out the co-occurring words and where the list of collocates should be cut-off. The usual way to obtain a collocational profile from a concordance is to scan within the span of 5 words to the right and to the left from the node. This method was not used here for several reasons: a) the resulting list of collocates does not reveal the relationship between the node and the collocate, b) it yields some noise as in the case of CD-ROM of English collocations, i.e. gives words as collocates which crop up within the span of 5 words but belong to a different clause than the node word (e.g. there are quite a few disconnected cases of the use of *"no, much, well"* under the top 20 collocates: *No further, mind*).

Thus we chose to move in our study step by step to the left due to a greater consistency of patterning to the left of the node word than to the right (cl. Sinclair 1996 B:84). A suitable cut-off point was made after the top twenty collocates in L1 position, i.e. the first word to the left from the node. In our case it consisted of 4/5 of all the concordance. The list of the top twenty L1 collocates is as follows:

Table 1. Top 20 collocates of *mind*

| Collocate in L1 | Frequency | Collocate in L1 | Frequency |
|---|---|---|---|
| *in* | 8462 | *its* | 574 |

| | | | |
|---|---|---|---|
| *my* | 7079 | *open* | 526 |
| *his* | 6879 | *their* | 506 |
| *the* | 5382 | *human* | 451 |
| *of* | 4178 | *never* | 348 |
| *your* | 3943 | *conscious* | 280 |
| *her* | 3076 | *creative* | 273 |
| *a* | 918 | *our* | 240 |
| *own* | 784 | *subconscious* | 221 |
| *and* | 744 | *public* | 204 |

After splitting the concordance into 20 files, each file was dealt with separately and split further into constituing files according to collocates in L2 or in some cases even L2 and L3 positions depending on the distance of the nearest notional word connected predicatively to *mind*. Concordance lines with no significant collocate in L2 or L3 positions were put into the file of *alia*, treated individually and classified on the basis of their syntactic pattern. Table 2 contains a further subdivision of collocates in L1 position from *my mind* and a wider context as in the case of *in my mind*.

Table 2. Subdivision of collocates in L1 position from *my mind*

1.1 change 1.1.1 change 1.1.2 changed 1.1.3. changing

1.2  cross 1.2.1. cross 1.2.2. crossed 1.2.3 crosses 1.2.4 crossing

1.3  in*

1.4 into 1.4.1 verbs of motion: come, creep, drill, drop, enter, flash, float, rush, etc.

1.5 keep 1.5.1 keep 1.5.2 keeping 1.5.3 keeps 1.5.4 kept

1.6 of 1.6.1 back 1.6.2  verbs of disappearance or removal: block, blot out, disappear, push, slip  1.6.3. out 1.6.3.1. am, is, was, were 1.6.3.2 verbs like: bored, scared, stoned, etc. 1.6.4 different nouns: cinema, balance, contents, eye, fabric, power, state, etc.

1.7 on 1.7.1 is, are,   was, were 1.7.2 have, has, had 1.7.3 verbs like: prey, stay, weigh, etc.

1.8 through 1.8.1 verbs of motion: flash, flit, go, pass, race, ran, stream, etc.

1.9 to 1.9.1 verbs of motion: bring, came, rush 1.9.2 adverb and conjunctions: and although, because, but, so, which, etc.

1.10 up 1.10.1 make, made

1.11 from 1.11.1 verbs of removal: ban, banish, block, clear, dismiss, dislodge, drive, erase, shift, etc. 1.11.2 far, further, remote, etc.

1.12 alia

---------------------------------------------------------------------------------------------------

-------------------

    * Since there was a great variety of collocates in L2 and further positions to the left, the whole group was subdivided into 10 subgroups according the syntactic patterns: 1.3.1 There + to be + something + in my mind

    1.3.2 *I* + verb of action + *in my mind*

    1.3.3 Something + verb of action + *in my mind*

    1.3.4 *It* + *is, was* + *in my mind* + to do something

    1.3.5 *I* + *have, had* + something + *in my mind*

1.3.6  *I* + verb of action + something + with/without + something + *in my mind*
1.3.7  Personal pronoun + *to be* + adjective or participle II + *in my mind*
1.3.8  Something + verb of motion + *in my mind*
1.3.9  *in my mind* is used parenthetically
1.3.10 *I* + verb of vision or creation + *in my mind's eye*

The outcome of this analysis was a great number of files compiled on the basis of a) outstanding collocations, b) most prominent syntactic patterns in case of the absence of significant collocations. The next step was to single out the collocations with *mind* on the basis of comparison and intuition so that a collocation would make a meaningful and partly autonomous word combination or a phrase. It was done manually since even partial automation in this step of analysis is impossible. Then it was possible to abstract information from the files and to join the files in relevant groups according to the meaning of collocates and/or the syntactic type of their relationship. The final step was to present the results of analysis.

While synthesising the prototypical pattern we applied different levels of abstraction. The most usual procedure was to lemmatise the significant collocates and abstract from their colligational pattern and specific grammatical forms, e.g. we joined separate files with *bear, bears, bearing, bore, born* under one lemma. Somewhat higher level of abstraction was to join different but still significant lemmas under the same syntactic pattern, e.g. *bear, keep, have something in mind*. Highest level of abstraction was used in those cases where a considerable number of collocates were joined as members of a semantic group under one common notion like 'motion', e.g. *Somebody's mind moves somewhere*  or even higher like 'action', e.g. *A person may act with something in mind.*

With the final purpose in mind, to present as detailed picture of usage as possible, we chose an elaborated version of the third type of lexicographic presentation, i.e. a corpus-based dictionary entry.  Our version differs from  the explanatory dictionary presentation in that it a) joins some similar phrases or includes them into more abstract patterns of usage as pointed above b) fills in the slots of more distant syntactic patterns with lists of collocates like nouns or noun phrases for subject and object positions, cl. objects in group 9 (see the pattern below). Names of more distant syntactic partners, such as objects, subjects,  are usually ignored by most dictionaries. Nevertheless they are relevant for the analysis of, for example, metaphorical usage of the noun and further processing of the pattern as well as for the comprehension of the concept.

The pattern of usage, presented below, comprises both sentences and phrases consisting of collocates of varying degree of co-occurrence. Those collocates with the total frequency in the span of 5 positions to both sides from the node word  ranges from the highest frequency down to 1000 times are given in underlined bold, e.g. **keep** one's mind **on** something; collocates ranging from 1000 to 5000 times of co-occurrence are given in italicised bold, e.g. *an open* mind; those ranging from 500 to 100 are given in simple bold, e.g. somebody's mind **races**; the remaining collocates are not highlighted. Words, added to generalise the whole group of collocates are simply underlined, in order to differentiate them from words, actually found in the concordance, e.g. somebody's mind moves somewhere. Here also belong *somebody*

and *something* as they stand for more distant syntactic partners with their specific manifestations exemplified in brackets.

## The pattern of usage of *mind*

### Sentences

1.  A person's mind <u>moves</u> (**races, runs, wanders,** goes round, reels, drifts, roams, vaults, flows, speeds, sinks, travels, spins, shifts back, rushes back, jumps over, leapfrogs ahead, tilts towards, slips into, slides away, sails, paces back, proceeds, circles etc.) somewhere or <u>does</u> <u>something</u> (**boggles, flashes, works,** suffers, recalls, plays, controls, embraces, shrinks, seizes, breaks, says, etc.) actively.

2.  A person's mind **is** of a certain <u>quality</u> or in a certain <u>state</u> (*open* **blank, clear, closed, fixed, focused, fresh, relaxed, full,** empty, agile, in turmoil, clouded, filled, fuzzy, intact, occupied, excited, blurred, shaken, overwhelmed, in paralysis, blown, on idle, concentrated, active, busy, scarred, screwed up, destroyed, fertile, acute, weary, sharp, attuned, stripped naked, feverish, fogged with fatigue, agitated, calm, numbed, dazzled, awake, trapped, chained, etc.).

3.  A person may **have, keep, put,** as well as **run, settle, set, fix** his/her mind **on** or **take** it **off** <u>something</u> (other things, marital problems, work, plan, business, something else, match, game, combat, new project, young ladies, tourism, task, other matters, politics, job, trouble, rheumatism, troubles, worries, weight, sex, etc. ).

3.1 A person may **put** his mind **to** something

4.  A person may **have, keep** his/her/a mind of a certain <u>quality</u> or in a certain <u>state</u> (*open,* **clear, focused, fresh,** active, alert, concentrated, dynamic, happy, sharp, centred, busy, at rest, occupied, steady, stable, balanced, turned to easy living, etc.)

5.  A person may <u>affect</u> **his/her mind** in various ways (**change, make up,** *open,* **boggle, cast, clear**, **close, control, fix, focus, free, read**, **relax, speak, turn,** adjust, affect, align, alter, balance, blank, blow, bring back, broaden, brush, build, calm down, challenge, cloud, concentrate, destroy, develop, distract, divert, divide, divorce, direct, dull, ease, educate, elevate, empty, exhaust, exercise, expand, explore, feed, fill, force, haunt, improve, invade, liberate, lose, model, occupy, penetrate, probe, project, poison, pollute, purify, push, reveal, shake, sharpen, shut down, soothe, stimulate, stretch, train, trouble, unravel, use).

5.1 A person, group of persons, an institution or a country may <u>affect</u> (**change,** etc.) his, her, their, its mind.

6.  A person may **be of/be with/have** (*half, a good*) a mind or **have it in** a mind <u>to do something</u> (accept challenges, argue otherwise, be realistic, escape, seize this opportunity, try anything, etc.)

6.1 A person may **be of/have** a mind of his/her own or *know* his/her own mind.

7.  A person may <u>act</u> (agree, create, determine, progress, think, decide, design, shop, go home, consider, train, carry out a plan, buy, examine, accept the offer, frame the budget, build, feel calmer, write, etc.) **with** <u>somebody or something</u> (audience, career, children, considerations, the World Cup, experience, factors,

family, goal, idea, image, needs, objective, Olympics, people, plan, possibility, reader, problem, project, purpose, question, safety, target, one thing, topic, thought, etc.) **in** mind.

8.     A person may **bear, have, keep** something (questions, doubt, point, work, idea, place, plans, response, task, certain things, statement, etc.) **in** mind.

8.1 A person may **have** something or something (see above) may **be in the** mind of a person.

9.     A person may **be** of a certain state or in a certain quality (**clear**, **relaxed,** certain, confident, disturbed, independent, aware, etc.) **in** his/her mind.

10.    A person may do something (go over it again, replay it, drive the race, carry a picture, go back, re-live, see, create, feel, realise, build a picture, etc.) **in** his/her mind.

11.    Something happens **in** somebody's mind (**images stick,** certain things flicker, words echo, points get firmly fixed, words get implanted, solution shape itself, questions linger, memories stand out, dream blossoms, figures blur, etc.)

12.    Something (**thought, fact, idea, it, saying, image, words, picture, question,** impulse, solution phrase, thing, **memories,**)  may move (***come, cross***, ***get, go***, **enter, flash, race, ran,**   creep, drill, float, pop, rush, slip, spring, flit, pass, stream, stray, crowd, flood, flow, leap, speed,  surge, parade, rush, penetrate, swim, whirl, sail, etc.) in different directions (**around, from, on, into, out, to**  one's mind) inside or inside-outside one's mind.

13.    A person may **have** something or do something (put, nag, keep, store, push, put, bring, pull, etc.) or something may **be** or happen in different parts (**back, stage,** cellar, centre, corner, compartment, depth, edge, forefront, front, recess, room, corridor, labyrinth, periphery, portion, section, screen, sky, surface, top, bottom, cavern, tunnels etc.)  **of** his/her mind**.**

14.    A person may *get (**put,** **cast, clear,** blot, push, block, shut, erase, ban, dislodge, banish, dismiss, remove, drive, shake, wipe, lift, expunge, elicit, vanish, wash, etc.) something (**thought, picture, idea,** pain, optimism, vision, burden, dangers, feelings of guilt, incident, etc.) **(out) of/ from** his/her mind.

15.    Something may **be far/get (out) of** (**get** away, slip, disappear, fade, etc.) **/from** his/her mind.

16.    A person may **be** ( be bored, stoned, scared, driven, gone, worried, etc.) **out of** his/her mind **with** something (worry, jealousy, frustration, etc.)

17.    Something (freedom, question, facts, pressure, last thing, other concerns, failure, problem, shame, sex, load, revenge, crime, death, demands, revolution, etc.) **is** (preys, weighs, stays, plays, etc**.) on** somebody's mind.

17.1 Something is a load  or a weight **off** somebody's mind.

18.    **There is** something (dates, items, memories, thoughts, questions, dreams, chance, confusion, decision, doubt, question, fear, images, way, failure, peace, picture, question, incident, loss, etc.) **in** one's mind.
       18.1 **It is/was in** somebody's mind to do something.

## Phrases

1.  Mind's something **(eye,** ear, capacity, **workings**, etc.).

2.  Something (**body,** feelings, eye, heart, language, memory, music, senses, skills, **spirit**, nervous system, material world, soul, life, hand, health ) and mind.

3. Something (cinema, balance, contents, product, structure, potency, fabric, **part, power**, side, **state, presence, <u>change,</u>** *frame,* habit, attitude, cast, *peace,* presence, levels, layers, windmills, blackboard, ills, muscles, **workings**, etc.) of mind.
4. *Human*, *good, right, open,* clear, closed, creative, conscious, fixed, fresh, political, public, racing, subconscious, sharp, strong, unconscious, young, sound, rational, collective, western, popular, brilliant, same, great, healthy, inquiring, dreaming, quick, dirty, troubled, logical, criminal, little, inner, analytical, legal, bad, scientific, inquiring, original, lively, fertile, twisted, limiting, modern, fine, sick, unsound, simple, suspicious, flexible, military, individual, interesting, active, strong, universal, different, trained, powerful, restless, independent, warped, fit, playing, waking, thinking, tired, narrow, questioning, critical, keen, tortured, agile, disordered, alert, normal, curious, easy, positive, similar, European, Soviet, inventive, ordinary, pure, calculating, remarkable, penetrating, divine, unquiet, romantic, youthful, empty, old, better, particular, searching, inquisitive, divided, average, weary, certain, real, practical, disturbed, extraordinary, peaceful, adult, musical, funny, directing, religious, liberal, bicameral, cultivated, free, understanding, blank, quiet, confused, small, big, growing, tidy, disciplined, incisive, tiny, official, deep, evil, prepared, female, negative, academic, older, weak, Australian, African, formidable, diseased, medieval, speculative, supreme, concentrated, objective, calm, focused, etc.

## Concluding remarks

It has to be emphasised that such a pattern of usage is not self-sufficient. It presents a stage of analysis between a concordance and a more elaborated description such as an entry in a general or a specialised dictionary, analysis of the semantic structure of a lexical unit, metaphorical concept, etc. The pattern of *mind* was designed so that it could serve as a raw material for further analysis and/or processing. This is only an attempt to present a usage pattern from a concordance that would include a wider context in the form of both syntactic patterns and lists of collocates. The ultimate goal is to make a complex lexical unit become evident.

No comment has been made on the meaning of *mind.* Thus the next step would be to relate uses to meanings, a task in which corpus analysis in the form of a usage pattern will inevitably play a central role revealing relevant features of the linguistic environment. In this case we can speak about semantic features, i.e. semantic preference and/or semantic prosody of a certain collocation. A manifestation of semantic preference is found in model 15 of the pattern above, e.g. *mind*, if preceded by directional adverbs like *to, into, out, from*, is used with the verbs of motion. An example of semantic prosody: collocation *on mind* is used with verbs *is, weighs, stays, plays, presses, preys* in L2 position and s great variety of nouns as subjects: *crash, burglary, eating, sex, topic,* etc. A closer look at the list of immediate and more distant collocates allows to postulate a semantic prosody of *worry* or *concern*. Both semantic preference and semantic prosody manifests itself in a more distant context of the core of a lexical item therefore it is important to include lists of most frequent contextual partners into the pattern of usage.

## Acknowledgements

## References

Altenberg, B. Amplifier Collocations in Spoken English // *English Computer Corpora: Selected Papers and Research Guide* / Ed. S.Johansson and A.-B. Stenstrom. Berlin: Mouton. 1991. 127-147.

Biber, D. Co-occurrence Patterns among Collocations: A Tool for Corpus-based Lexical Knowledge Acquisition // *Computational Linguistics*. 1993. **19**. 549-556.

Biber, D. *Investigating Language use through Corpus-based Analyses of Association Patterns // International Journal of Corpus Linguistics*. 1996. **1** (2). 171-198.

Church, K., Hanks, P. Word Association Norms, Mutual Information, and Lexicography // *Computational Linguistics*. 1990. **16**. 22-29.

Hanks, P. Contextual Dependency and Lexical Sets // *International Journal of Corpus Linguistics*. 1996. **1** (1). 75-98.

Hanks, P. Enthusiasm and Condescension // *ACTES EURALEX'98. PROCEEDINGS. Papers submitted to the Eight EURALEX International Congress on Lexicography in Liege, Belgium. University of Liege, English and Dutch Departments*. 1998. 151-166.

Levin, B., Song, G., and Atkins, B.T.S. Making Sense of Corpus Data: A Case Data of Verbs of Sound // *International Journal of Corpus Linguistics*. 1997. **2** (1). 23-64.

Louw, W.E. Irony in the text or insincerity in the write? The diagnostic potential of semantic prosodies // *Text and Technology* / Ed. M. Baker et al. John Benjamins, Amsterdam. 1993. 152-176.

Marcinkevičienė, R. Atminties labirintuose. Kognityvinės ir tekstynų lingvistikos sąveika // *Darbai ir Dienos*. 1999. **10** (19). 109-124.

Marcinkevičienė, R. Kolokacija: tyrimo objektas, aspektas, metodai // *Lituanistica*. 1995. **2** (22). 40-54.

Sinclair, J. The Empty Lexicon // *International Journal of Corpus Linguistics*. 1996 A. **1** (1). 99-120.

Sinclair, J. The Search for Units of Meaning // *Textus*. 1996 B. **IX** . 75-106.

Sinclair, J. The Lexical Item // *Contrastive Lexical Semantics*. Series: *Current Issues in Linguistic Theory*. Amsterdam/Philadelphia: John Benjamins. 1998. ….

Sinclair, J. The Computer, the Corpus and the Theory of Language // *Lingua*. 1999. **1.** 24-32.

Stubbs, M. Collocations and Semantic Profiles: On the Cause of the Trouble with Quantitative Studies // *Functions of Language*. 1995. **2.** 23-55.

Teubert, W. Corpus Linguistics - A Partisan View // *TELRI (Trans-European Language Resources Infrastructure: Concerted Action in the Framework of the Copernicus Program)*. 1999. **8.**. 4-19.

Williams, G.C. Collocational Networks: Interlocking Patterns of Lexis in a Corpus of Plant Biology Research Articles // *International Journal of Corpus Linguistics*. 1998. **3** (1). 151-172.

**Lexicographic sources**

*The Bank of English* http://titania.cobuild.collins.co.uk

Benson, M., Benson, E., Ilson, R. *The BBI Combinatory Dictionary of English: A guide to Work Combinations.* Amsterdam/Philadelphia: John Benjamins Publishing Company, 1986.

*Collins Cobuild English Dictionary.* HarperCollins Publishers, 1995.

*Collins Cobuild English Collocations on CD-ROM. A Comprehensive Database of Common Word Patterns from the Bank of English.* HarperCollins Publishers, 1995.

*Longman Dictionary of Contemporary English.* Third edition. Longman dictionaries, 1995.

## ŽODŽIO JUNGLUMO MODELIS TEKSTYNŲ LINGVISTIKOS ASPEKTU

Rūta Marcinkevičienė

Reziumė

Darbas remiasi teorine prielaida, jog reikšmė randasi platesniame nei atskiro žodžio kontekste, todėl jai tirti labai svarbus yra žodžio junglumo modelis. Šį modelį galima parengti analizuojant iš tekstyno gautą tiriamojo žodžio konkordansą. Čia aptariami keli aprašomieji ir žodyniniai žodžių bei jų junginių vartosenos modeliai ir siūlomas savas variantas, užimantis tarpinę vietą tarp konkordanso ir žodyninio straipsnio ar kokio kito tolesnės analizės rezultato. Jis iliustruojamas anglų kalbos daiktavardžio *mind* vienaskaitos vartosenos aprašu, pusiau automatiškai gautu iš daugiau nei 50 000 konkordanso eilučių, ir nuoseklia konkordanso analizės bei junglumo modelio sintezės eiga.

Vytauto Didžiojo universiteto                          Įteikta

Kompiuterinės lingvistikos centras                          2000 m. sausio 10 d.