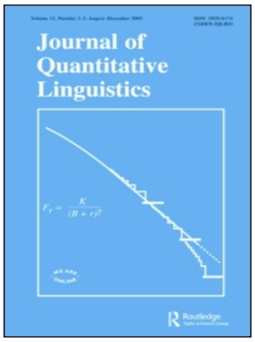

Journal of Quantitative Linguistics

ISSN: 0929-6174 (Print) 1744-5035 (Online) Journal homepage: http://www.tandfonline.com/loi/njql20

Semantic Fields: Formal Modelling and Interlanguage Comparison

Karina Dullieva

To cite this article: Karina Dullieva (2017) Semantic Fields: Formal Modelling and Interlanguage Comparison, Journal of Quantitative Linguistics, 24:1, 1-15, DOI: 10.1080/09296174.2016.1239400

To link to this article: https://doi.org/10.1080/09296174.2016.1239400

Published online: 24 Nov 2016.

JOURNAL OF QUANTITATIVE LINGUISTICS, 2017
VOL. 24, NO. 1, 1–15
http://dx.doi.org/10.1080/09296174.2016.1239400

Semantic Fields: Formal Modelling and Interlanguage Comparison

Karina Dullieva

Intercultural Communication Department, Kazan Federal University, Kazan, Russia

ABSTRACT
The research deals with a semantic field constructing method and comparative analysis of two semantic fields. Semantic features of lexemes are on the basis of creating an abstract model of concept that semantically unites all the units of compared semantic fields. The method of contrasting semantic fields is on the basis of "abstractness-concreteness", which can give quantitative estimate of differences between semantic fields. Cooking verbs semantic fields are taken into consideration. It is shown that the semantic field of cooking verbs in English is more specific compared with that in Russian.

1. Introduction

To study structural relations between words in the lexical system of the language there are different word classifications: parts of speech, word-forming nests, synonym sets, antonym pairs, semantic fields and other associations. The most difficult to construct is a semantic field. This is due to the fact that at present formalism for describing semantic fields is poorly developed.

According to the semantic fields theory founder J. Trier (see Trier, 1931), the semantic field is a group of words which are united according to a common basic semantic component. (A detailed description of modern semantic fields theory is given in Vassilyev (1974), Geeraerts (2010), a brief history of its development is in Kleparski and Rusinek (2007)). At the same time, the words within the semantic field may be interconnected in a variety of semantic relations, for example, hyponymy, hyperonimy, meronimy etc. (see Cruze, 1986; Geeraerts, 2010). These relations are mainly due to the semantic structures of words in semantic fields. The meaning of each word is made up of elementary semantic units, attributes or semes (see Novikov, 1982; Popova and Sternin, 2009). It is important to understand that these semes may also be linked by

CONTACT Karina Dullieva dulliev@yahoo.com

some semantic relations. In our opinion, these relations should determine not only the semantic structure of each word but also the whole semantic field.

Usually, the semantic field structure is described as a tree diagram with one root (see Lehrer, 1974; Lipka, 1992). The elements of the tree are the words taken from a semantic field, connected by hyperonimy–hyponimy relations. It is quite possible that such a detailed semantic field description is reasonable in certain cases; however, there are problems where it is not applicable. The article deals with the problem of interlanguage comparison of semantic fields having the common integral seme.

The difficulty faced by researchers when contrasting semantic fields is the necessity of taking into account features of semantic field structures and the organization of seme within sememes, in other words, the semantic fields contrasting should be carried out not only on some specific lexemes, but at the same time and on the structural features of the semantic field. The necessity of taking into account the semantic field structures leads to the problem of the semantic field structure concept explication. In this article, we have attempted to define the semantic field structure through the abstract model of the basic concept. The basic concept unites semantically all the units of the contrasted semantic fields and corresponds to the integral seme. The model is defined by formal mathematical constructions. Let's specify at once, the description is not the only one. There can be suggested descriptions based on binary relations IS-A and INSTANCE OF, special relations of pre-order, Boolean algebras and so on (see Tsalenko, 1989). There are also semantic fields descriptions in the form of hypergraphs, which are built with sets of first t-images (see Dyvik, 2005) or with the help of numerical characteristics of word meanings proximity (see Streib, 2014).

Once the structure of the contrasted semantic fields has been built, one can try singling out a particular structural feature according to which one can make contrasting. It is desirable the structural feature should allow a quantitative description; for example, when it characterizes the "position" of a lexeme in the semantic field. As such a feature, we selected a sign that reflects the level of abstractness or concreteness of the lexeme within the frame of the semantic field. This feature is gradual, so we can graphically depict its value on a straight line where the points represent the lexeme in accordance with the feature value.

The semantic fields constructing procedure with their further contrasting can be divided into several steps of the study. Chronologically, the implementation of each step requires the results of the previous ones.

Step 1. Selecting lexemes belonging to a semantic field out of dictionaries in concerned languages.

Step 2. Singling out semes out of lexeme values. (Component analysis, see Gak (1977); Lipka (1992); Popova and Sternin (2009); Geeraerts (2010)).

Step 3. Creating an abstract model of the basic concept (semantic field structures) in accordance with the results of Step 2.

Step 4. Formal representation of lexemes in accordance with the built basic concept model.

The realization of Steps 1–2 is more or less ordinary and does not cause any difficulties. In the following sections we will have a detailed look at Steps 3–5.

Step 5. Semantic fields contrasting on the basis of 'abstractness -concreteness'.

The implementation of the Steps 1–2 is more or less standard and usually does not cause any difficulties. They must meet the conditions briefly listed in Section 2. Section 3 is devoted to a detailed description of Steps 3–4, and Section 4 describes Step 5. Section 5 deals with the application of the proposed methods, where it is considered the cooking verbs semantic fields contrasting in Russian and English languages on the basis of "abstractness–concreteness".

2. Preliminaries

Throughout this paper, the semantic field implies a group of words, which are united by a common basic semantic component (see Lipka, 1992; Popova and Sternin, 2009), normally referred to as an integral seme or a hyperseme. We suppose that contrasted semantic fields have one and the same integral seme.

Semantic differences of words within a semantic field are characterized by differential semes or hyposemes. Some lexemes may have several similar or different differential semes. All these semes are linked to each other by different relations and taken together form the semantic field semantic structure.

While implementing Step 1 one should take into consideration lexemes polysemy. In this article, it is done as follows. It is supposed that lexico-semantical variants of a word (see Novikov, 1982; Popova and Sternin, 2009) are considered to be various elements of the semantic field. To cut a long story short, let us take it for granted to refer to lexico-semantical variants of words as lexemes. While studying semantic fields consisting of verbs, the latter should be in the infinitive form and in the imperfective aspect. Verbs having one and the same root are considered to be forms of one and the same unit.

The implementation of Step 2 should satisfy the following conditions:

1. The lexeme takes into account only denotative macro-component of lexeme values, i.e. that part of the lexeme lexical meaning, which corresponds to a particular object designated by a word. This value macro-component reflects, first of all, a natural ontological division of objects, attributes, properties, actions, processes, events and conditions.
2. The set U of all semes taken into consideration includes only the integral and differential semes of all investigated semantic fields. The process of selecting semes is carried out "downward", beginning with the integral seme by sequential specifying semes.

3. Specification of each seme means finding out its differentiating parameters, hereinafter referred to as seme variables.
4. Every seme is the value of a seme variable.
5. The semes that refer to one and the same seme variable can be connected only by disjunction, and the semes corresponding to different seme variables can be connected only by conjunction.

Furthermore, if the seme u is differentiated by the seme variables $t_1, \ldots, t_k$, there is a variable t_j is complex, composite, i.e. it is differentiated by the variables $t_{j1}, \ldots, t_{jm}$, then instead of the set of the variables $t_1, \ldots, t_k$ for the seme we will conseder the set $t_1, \ldots, t_{j1}, \ldots, t_{jm}, \ldots, t_k$, considering the sets to be equivalent. With regard to the agreement, everywhere below we believe that the process of formalizing the model of the basic concept involves only simple parameters.

3. Formal models of basic concepts (semantic field structure) and a lexeme

While describing the abstract model of the basic concept we take into account the conditions mentioned in the previous section.

The model of the basic concept is schematically represented as a hierarchy with a single "root", where there are both simple tops, differentiated on no more than one parameter, and complex ones, differentiated on several parameters. Moving from "the root" down to "branches", we get new concepts derived from the basic one. One can move from the complex tops along one or more branches at the same time. As one can see below, all derived concepts that result from such movements in the hierarchy can be described recursively using the mathematical language of the theory of relations. The advantage of this approach is that, firstly, it reflects a functional nature of human natural languages, and, secondly, it allows us to submit into consideration a multilevel model of the basic concept in a compact and comprehensible form. Let us move on to the precise formulations.

Let U be the set of all the semes singled out in Step 2. The elements of the set will be referred to as vertices. Each vertex $u \in U$ is assigned a set $\varphi(u) \subset \{u\} \times \underbrace{U \times \ldots \times U}_{k(u)}$, where $k(u) \in \{0, 1, \ldots\}$. The sequence number, starting with the second, of each factor in $\varphi(u)$ corresponds univocally to a seme variable differentiating u, which has exactly $k(u)$ seme variables. The set $\mathrm{Pr}_i\, \varphi(u)$, $i \in \{2, \ldots, k(u)+1\}$, is a set of values of i-th seme variable. Further on $\mathrm{Pr}_i\, \varphi(u)$ denotes i-projection of the set $\varphi(u)$, that is

$$\mathrm{Pr}_i \varphi(u) = \{v \in U : (\{u\} \times \underbrace{U \times \ldots \times \overbrace{\{v\}}^{i} \times \ldots \times U}_{k(u)}) \cap \varphi(u) \neq \emptyset\}.$$

We assume that the set $S = \{\varphi(u){:}u \in U\}$, which is associated with an abstract model of the basic concept and hereinafter referred to as a semantic tree with a set of vertices U, satisfies the following axioms.

S1. $\exists! u^0 \in U\ \forall u \in U\ \forall i \in \{2, \ldots, k(u)+1\}{:}u^0 \notin \mathrm{Pr}_i\, \varphi(u)$.

S2. $\cup\{\mathrm{Pr}_i\, \varphi(u){:}u \in U, i \in \{2, \ldots, k(u)+1\}\} = U \backslash \{u^0\}$.

S3. $\forall u, v \in U\ \forall i \in \{2, \ldots, k(u)+1\}\ \forall j \in \{2, \ldots, k(v)+1\}{:}\ \mathrm{Pr}_i\, \varphi(u) \cap \mathrm{Pr}_j\, \varphi(v) = \emptyset$.

Let us comment briefly on the above-mentioned axioms. The first one postulates the existence of the only root vertex u^0 of the semantic tree S: the vertex cannot represent any value of a seme variable of $u \in U$. In contrast, according to the second axiom, any other vertex distinct from u^0, is necessarily the value of a seme variable for some vertex $u \in U$. It is obvious that u^0 corresponds to the integral seme of the semantic fields in question. The vertices $u^{\mathrm{f}} \in U$ having an empty set of parameters: $k(u^{\mathrm{f}}) = 0$, will be called final vertices. The last two axioms state that the subset $U \backslash \{u^0\}$ is broken by the relationships $\varphi(u)$, $u \in U$ in disjoint subsets $\mathrm{Pr}_i\, \varphi(u)$, $i \geq 2$.

Let us describe the rules for constructing formal expressions, hereinafter referred to as terms, which will describe concepts derived from the base one within the semantic tree S, as well as lexemes of the investigated semantic fields, because each lexeme corresponds unambiguously to a concept. The definition is given recurrently.

1. Every set $\{u\} \subset U$ is a term (atomic).
2. For each vertex $u \in U$, any non-empty set $a \subset \varphi(u)$ is a term; for the sake of convenience, these terms will be written in the form $u(\mathrm{Pr}_2\, a, \ldots, \mathrm{Pr}_{k(u)+1}\, a)$.
3. If a_1 and a_2 are terms, so that $\exists i \in \{2, \ldots, k(\mathrm{Pr}_1\, a_1) + 1\}{:}\ \mathrm{Pr}_1\, a_2 \subset \mathrm{Pr}_i\, a_1$, then $\mathrm{Pr}_1\, a_1 \times \ldots \times \left((\mathrm{Pr}_i\, a_1 \backslash \mathrm{Pr}_1\, a_2) \cup a_2\right) \times \ldots \times \mathrm{Pr}_{k(\mathrm{Pr}_1\, a_1)+1}\, a_1$ is a term. It is obvious that this term may be also represented in the form of $S_1 \cup S_2$, where $S_1 = \mathrm{Pr}_1\, a_1 \times \ldots \times (\mathrm{Pr}_i\, a_1 \backslash \mathrm{Pr}_1\, a_2) \times \ldots \times \mathrm{Pr}_{k(\mathrm{Pr}_1\, a_1)+1}\, a_1$, $S_2 = \mathrm{Pr}_1\, a_1 \times \ldots \times a_2 \times \ldots \times \mathrm{Pr}_{k(\mathrm{Pr}_1\, a_1)+1}\, a_1$.
4. The subseet $a = \bigcup\limits_{\alpha} S_\alpha$, $S_\alpha \subset U \times \ldots \times U$, a term, if and only if, it can be obtained from the conditions 1–3.

In the condition 3, as well as in 2, the terms will be recorded in the form $\mathrm{Pr}_1\, a_1(\mathrm{Pr}_2\, a_1, \ldots, \left((\mathrm{Pr}_i\, a_1 \backslash \mathrm{Pr}_1\, a_2) \cup a_2\right), \ldots, \mathrm{Pr}_{k(\mathrm{Pr}_1\, a_1)+1}\, a_1)$ and in 1 they will be written in the form of $u(\ )$, or, in short, u. In addition, the terms u and $u(\mathrm{Pr}_2\, \varphi(u), \ldots, \mathrm{Pr}_{k(u)+1}\, \varphi(u))$ where $u \in U$ will be identified.

Every term $l = \bigcup\limits_{\alpha} S_\alpha$, obtained by using rules 1–4, so that $\forall \alpha\ \mathrm{Pr}_1\, S_\alpha = \{u^0\}$, is associated with a concept that is induced from the semantic tree S. The procedure of constructing terms and conditions S1–S3 also implies that

$$\forall l\ \forall \alpha\ \forall i \geq 2\ \exists! u \in U\ \exists! j \in \{2, \ldots, k(u)\}{:}\mathrm{Pr}_i S_\alpha \subset \mathrm{Pr}_j \varphi(u), \qquad (*)$$

i.e. for any α and any, starting from the second, i-th projection S_α of an arbitrary term l there is only one pair of "vertex (seme) u - its j-th seme variable" so that this projection is the subset of j-th seme variable.

As noted above, every lexeme corresponds univocally to a concept induced by the semantic tree S, therefore, there is the term l corresponding to every lexeme out of the semantic field.

Summarizing the formal description, we find out that the abstract model of the basic concept that defines the structures of the investigated semantic fields, represents the semantic tree S and the semes set U, and the seme variables and their values (semes) are determined explicitly by the relationships $\varphi(u)$ (Step 3). Lexemes of the semantic fields are represented by the corresponding sets of terms l (Step 4).

4. Semantic fields contrasting

As mentioned in the Introduction, the article deals with inter-language semantic fields contrasting, that of the first and the second semantic fields, on the basis of "abstractness-concreteness". A quantitative description of this feature is carried out by the weight concept. Let us define it.

The weight of the lexeme $l = \bigcup_{\alpha} S_\alpha$ is a number $P(l)$ calculated by the formula

$$P(l) = \sum_{\alpha} \prod_{i \geq 2} \frac{|\operatorname{Pr}_i S_\alpha|}{N(S_\alpha, i)}, \qquad (**)$$

Here the symbol $|\ |$ denotes the set elements number; $N(S_\alpha, i) = |\operatorname{Pr}_j \varphi(u)|$, where u and j are determined by (*).

Using the formula (**), we can calculate the lexemes weights in the contrasted semantic fields. Furthermore, the formula (**) can be used to determine the weights of the concepts derived from the basic one.

It is obvious that the less the lexeme weight value is the more particular, specific, concrete, it is. All weight values are located on the half-interval (0,1], and the maximum weight which is equal to one, can be achieved only on the root node u^0. The weight equal to zero is never achieved, no matter what lexeme it is. One can say that lexeme weights are, in some sense, their position coordinates in the hierarchy of the basic concept.

The weight calculating formula (**) is as follows. When one moves down the basic concept model hierarchy, lexeme weights decrease multiplicatively, that is, the weight of every lexeme $l_i = \bigcup_{\alpha} S_\alpha$ differs from the weight of the higher lexeme $l_j = \bigcup_{\beta} S_\beta$ in the hierarchy (i.e. $\forall S_\alpha \forall S_\beta \ \bigcup_{q \geq 1} \operatorname{Pr}_q S_\beta \subset \bigcup_{q \geq 1} \operatorname{Pr}_q S_\alpha$) in a numerical factor lying in the interval (0,1). As a result of this, plenty of weights are concentrated near zero, and the corresponding distribution of the weights on the interval [0,1] according to the frequency of their occurrence is approximately to zero. Unfortunately, this form of weights distribution representation is not very convenient because it does not describe very clearly the type of weight changes. This visibility can be achieved by the transition from the multiplicative calculating the weights to the additive one using logarithms. The base of logarithm to find lexeme weights logarithms here is equal to 10. Let us note that nothing prevents us from choosing another number that is more than zero. Since a lexeme weight is less than one, and the logarithm base is more than one, it makes sense to multiply every logarithm by minus one, thereby making it a non-negative value.

The terms corresponding to the first contrasted semantic field are marked with the symbol l_i^1, where $i \in \{1, ..., m\}$, m equals to the lexeme numbers from the first semantic field. Similar notation is used for the second sematic field: l_j^2, $j \in \{1, ..., n\}$. Where our reasonings are applicable simultaneously to any lexemes regardless of the semantic field, the superscripts in the notations l_i^1, l_j^2 will be omitted. Besides, suppose $L^1 = \{l_1^1, ..., l_m^1\}$, $L^2 = \{l_1^2, ..., l_n^2\}$.

Thus, we see that the contrasted semantic fields are described with the couple sets $(l_i^1, I(l_i^1))$, $i \in \{1, \ldots, m\}$ and $(l_j^2, I(l_j^2))$, $j \in \{1, \ldots, n\}$, corrispondingly, where $I(l) = -\lg P(l)$. The set comparison can be done in different ways. In this paper, we compare some of their statistical characteristics.

Let us consider the mean arithmetic values $\overline{I^1}$ и $\overline{I^2}$ for values $I(l_i^1)$, $i \in \{1, \ldots, m\}$ and $I(l_j^2)$, $j \in \{1, \ldots, n\}$, correspondingly. By finding anti-logarithm of these values one can find the corresponding weight values $\overline{P}^1$ и $\overline{P}^2$. Suppose, that $\overline{P}^1 > \overline{P}^2$. Then this inequality admits the following interpretation: the second semantic field in comparison with the first one has a shift towards a more special meaning of words, and the difference is one seme variable, whose set of values is in average $\overline{P^1}/\overline{P^2}$ elements. If $\overline{P}^1 = \overline{P}^2$ then semantic fields shift according to each other in the direction of abstractness or concreteness in average does not exist.

In order to visualize the distribution of the numbers $I(l)$, the histograms of their relative frequencies are used. For its construction, we will limit the possible values set of the value $I(l)$ by the segment $\Delta = \left[\min\left\{\min_{1 \le i \le m} I(l_i^1), \min_{1 \le i \le n} I(l_i^2)\right\}, \max\left\{\max_{1 \le i \le m} I(l_i^1), \max_{1 \le i \le n} I(l_i^2)\right\}\right]$. Further, let us divide the segment Δ into $s = 1 + \lfloor \log_2 \max\{m, n\} \rfloor$ segments of equal length Δ_q, $q \in \{1, \ldots, s\}$, and on each of them we will build two rectangles with heights proportional to the respective relative frequencies m_q and n_q of falling values $I(l_i^1)$, $i \in \{1, \ldots, m\}$ and $I(l_j^2)$, $j \in \{1, \ldots, n\}$. The frequencies m_q and n_q for the first and the second semantic fields are calculated independently one from another: having different denominators – m and n, respectively; and the values $I(l)$, that fall on the bounds of two adjacent segments, are referred to the right one.

It is obvious that a set of numbers $I(l_i^1)$ and $I(l_j^2)$, where $i \in \{1, ..., m\}$, $j \in \{1, ..., n\}$ are distributed unevenly on the interval Δ. This means that on Δ there are the segments Δ_q with sufficiently different "densities" of numbers clusters $I(l)$, or in other words, with different rectangles heights over the segments Δ_q.

The comparison of the given rectangles on each segment Δ_q to quantify semantic fields semantic distinction matched on the basis of "abstractness–concreteness". Any segment Δ_q can be interpreted as follows.

Let a seme variable of each seme consists on an average of h elements. (The number h can be estimated, for example, by the value $\frac{|U \setminus \{u^0\}|}{\sum_{u \in U} k(u)}$). Then each segment Δ_q is characterized on an average by $c = \frac{10^{-|\Delta|/s}}{h}$ seme variables (where $|\Delta|$ is the length of the segment Δ). Hence we have c seme variables on the segment Δ_q. Thus, the lexemes taken from two adjacent segments differ on the average by almost c seme variables.

5. Application to contrasting cooking verbs semantic fields in the Russian and English languages

As an application of developed in Sections 2–4 methods, we consider the cooking verbs semantic fields contrasting in the Russian and English on the basis of "abstractness–concreteness".

In Step 1 (see Section 1), we singled out 71 cooking verbs in Russian and 109 cooking verbs in English (Appendix I). These verbs we extracted from bilingual modern, explanatory dictionary, synonymous dictionaries, thesaurus of the Russian and English languages (see Appendix II for a complete list of dictionaries). Illustrations of the listed dictionaries were used as well. All the verbs were taken in the infinitive form and in the imperfective aspect. Verbs having one and the same root were considered to be the forms of one and the same unit.

The results of the implementing Steps 2–4 and the lexemes weights values because of their large amount can be found in Dullieva (2015), where for the sake of clarity, the semantic tree of the base concept "cooking a meal" is represented in the form of a graphical scheme. In this article, we mention, as an example, only the data obtained for the lexeme "parboil": $(l^E_{parboil}, I(l^E_{parboil}))$, $l^E_{parboil} = T\left(T_1\left(p_1, r\left(r^1_1(r^1_{11}), R^2, R^3\right), u_6, v^1_{11}, V^1_2, v^1_{31}, v^2_2, v^3_1(v^3_{15}), v^4_1\right)\right)$, $I(l^E_{parboil}) = -\lg P(l^E_{parboil}) \approx 6.845$, $P(l^E_{parboil}) = \frac{1}{3} \cdot \frac{1}{2} \cdot \frac{1}{10} \cdot \frac{1}{2} \cdot \frac{1}{9} \cdot \frac{1}{3} \cdot 1 \cdot 1 \cdot \frac{1}{9} \cdot \frac{1}{2} \cdot 1 \cdot \frac{1}{2} \cdot \frac{1}{3} \cdot \frac{1}{2} \cdot \frac{1}{5} \cdot \frac{1}{2} \approx 1.43 \times 10^{-7}$. (The interpretation of the symbols used in the recording $l^E_{parboil}$, as well as that of the symbols used while recording other lexemes can be found in Dullieva (2015).)

Direct calculations show that, mean values for $I(l^R_i)$, $I(l^E_i)$ (R – Russian, E – English) $\overline{I^R} = 3.605$ и $\overline{I^E} = 3.994$ up to the fourth decimal place. The corresponding weights values $\overline{P^R} = 0.000248$ и $\overline{P^E} = 0.000101$ up to the 7th decimal place. Thus, the English semantic field compared to the Russian one has a shift towards a more special meaning of the words, the difference is one seme variable, the value set of which is $\overline{P^R} \big/ \overline{P^E} \approx 2.5$ elements on an average.

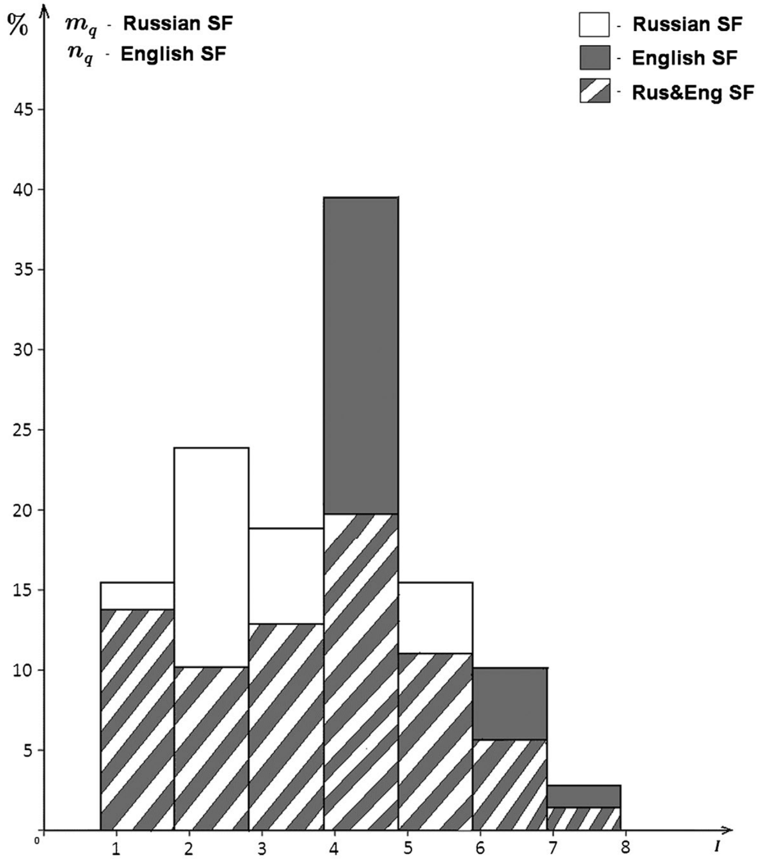

Figure 1. Combined histograms for the distribution of the value *I* for cooking verbs semantic fields in Russian and English (SF, semantic field).

The results of grouping the lexemes on the segments Δ_q for the Russian and English semantic fields are given in Appendix I, and their combined histograms in Figure 1. It was also found that $\Delta \approx [0.778;7.925]$, $s = 7$, $h \approx 3.57$, $c = 2.93$.

From Figure 1 it is clear that, starting from the fourth grouping interval (hereinafter all the intervals are numbered in ascending order from the left to the right), the number of lexemes with a narrower value decreases monotonically to the right for both languages. The largest number of lexemes in the Russian and the English languages, respectively, fall on the 2nd and 4th segments. The relative frequencies for the English semantic field exceed approximately two times the corresponding relative frequencies for the Russian one on the 4th, 6th, 7th intervals. The ratio of the same frequencies on the intervals 1–3, 5 is about 0.6. Thus, the distribution of relative frequencies also shows that the English cooking verbs semantic field is more special compared to the Russian one. Particularly striking is the part of the histogram, which is located on the 4th interval: the height of the rectangle for the English language is 3.2 times more than the height of any other rectangle for the same language on an average.

6. Conclusion

Let us note some features of the proposed methods. Above all, from the formula (**) it follows that if any two lexemes $l_q = (\underset{\alpha}{\cup} S_\alpha) \cup S'$ and $l_r = (\underset{\alpha}{\cup} S_\alpha) \cup S''$ satisfy the condition $\exists j \, \forall i \neq j \; (\mathrm{Pr}_i\, S' = \mathrm{Pr}_i\, S'', U \ni u_q = \mathrm{Pr}_j\, S' \neq \mathrm{Pr}_j\, S'' = u_r \in U)$, i.e. l_q and l_r differ in one seme according to one seme variable, then $P(l_q) = P(l_r)$. Hence it follows that any two different semes, which are the values of one seme variable, make the same contribution in the weight calculation, or, in other words, the semes are "uniformly distributed" on the set of corresponding seme variable. Generally speaking, it is not always the case, for example, in cases where there is a large spread of seme extensionals: there are simultaneously large extensional semes and narrow extensional semes. However, our assumption of semes distribution in a uniform way is only an assumption and can be modified if necessary. Further, the main contrasting method advantage is the quantitative estimate of the "abstractness–concreteness" characteristic, which is achieved by introducing the lexeme weight concept. Though, it should be noticed that in the article in question the parameter "abstractness–concreteness" cannot be fully referred to the dichotomy "wide extensional–narrow extensional" in the sense Gak (1977). It is more exactly, it just shows how much the lexeme meaning is "saturated" with seme variables and their values (semes), i.e. it shows the degree of specialization of the lexeme meaning. The main difficulty of the method is associated with the construction of the basic concept semantic tree: it is made up with the help of lexeme semes received as a result of the componential analysis as well as taking into account the ontological division of objects, actions and processes represented by lexemes. In particular, it suggests updating the semantic tree when one changes the semantic field composition (introducing or eliminating lexemes).

Now let us consider briefly how one can carry out the methods modifications. The first direction has already been mentioned above and implies a transition from the uniform weight distribution to the distribution that takes into account semes extensionals. The second direction is to find a replacement to the "abstractness–concreteness" feature, which would also permit quantitative representation and which would take into account the semantic tree structure in a more particular way. For example, one can move to a two-dimensional parameter, one coordinate of which would reflect the lexeme seme variables position in the semantic tree hierarchy, and the other one would represent the lexeme semes quota in the seme composition of correspondent seme variables. In addition, while formalizing the basic concept model one can have a try to reject the semes variables simplicity requirement (see Section 2). Furthermore, the semantic distance coefficient can be taken as a quantitative factor, which was proposed by Titov (2002), or indices listed in Sternina (2014). The third direction is contrasting semantic fields in three and more languages. Yet, it should be borne in mind that the comparison of a large number of semantic fields can be rather problematic. Therefore, the implementation of various contrasting semantic fields steps listed in the introduction are to be carried out by means of automated computational procedures that allow the possibility of using them on a computer.

Disclosure statement

No potential conflict of interest was reported by the author.

References

Cruze, David A. (1986). *Lexical Semantics*. Cambridge: Cambridge University Press.

Dullieva, K. (2015). *Contrasting The Cooking Verbs in Russian and English*. Unpublished row data. URL Retrieved from http://shelly.kpfu.ru/e-ksu/docs/F1709401788/Contrasting.The.Cooking.Verbs.pdf.

Dyvik, Helge (2005). Translations as a semantic knowledge source. In Margit Langemets & Priit Penjam (Eds.), *Proceedings of the Second Baltic Conference on Human Language Technologies* (pp. 27–38). Tallinn: Tallinn University of Technology & Institute of the Estonian Language.

Gak, Vladimir G. (1977). *Comparative Lexicology Based on French and Russian Languages*. (Sopostavitelnaya lexikologiya Na materiale frantsuzskogo i russkogo yazykov). Russian: International Relations. Moskva.

Geeraerts, Dirk (2010). *Theories of Lexical Semantics*. Oxford: Oxford University Press.

Kleparski, Grzegorz A., & Rusinek, Angelina (2007). The tradition of field theory and the study of lexical semantic change. *Studia Anglica Resoviensia*, *4*(UR), 188–205.

Lehrer, Adrienne (1974). *Semantic Fields and Lexical Structure* Amsterdam: North-Holland.

Lipka, Leonhard (1992). *An Outline of English Lexicology: Lexical Structure, Word Semantics, and Word-Formation* Tübingen: Max Niemeyer Verlag.

Novikov, Lev A. (1982). *Semantics of the Russian Language*. Moskva: Vysshaya Shkola.

Popova, Z. D., Sternin I. A. (2009), *Lexical System of Language* (Lexicheskaya sistema yazyka: Vnutrennyaya organizatsiya, kategorialnyi apparat). (Russian) Moskva: Librkom.

Sternina M. A. (2014). *Sopostavitelno-parametricheskii metod lingvisticheskikh issledovanii* [Comparative-parametric method of linguistic research]. (Russian) Voronezh: Istoki.

Streib, Heinz (2014). *Components of the Semantic Field of "Spirituality" and "Religion" Emerging from Semantic Differentials*. Presented at the International Seminar on Religious Education and Values (ISREV), York, UK.

Titov, V. T. (2002). *Obshshaya kvantativnaya lexicologiya romanskich yazykov* [General quantitative lexicology of Roman languages]. (Russian) Voronezh: Izdatelstvo Voronezhskogo Universiteta.

Trier, Jost (1931). *Der deutsche Wortschatz im Sinnbezirk des Verstandes* [Consciousness semantic filed in German]. Heilderberg: Winter.

Tsalenko, Michael Sh. (1989), *Modelirovanie semantiki v bazakh dannykh* [Semantics Modelling in Databases]. (Russian) Problemy Iskusstvennogo Intellekta, 14. Moskva: Nauka.

Vassilyev, Leonid M. (1974). The theory of semantic fields: A survey. *Linguistics*, *137*, 79–94.

Appendix I. Results for grouping lexemes both in Russian and English cooking verbs sematic fields

Table 1. English lexemes.

i	1	2	3	4	5	6	7
Δ_q	[0.778; 1.799)	[1.799; 2.820)	[2.820; 3.840)	[3.840; 4.862)	[4.862; 5.882)	[5.882; 6.903)	[6.903; 7.925]
n_q	0.137614679	0.100917431	0.128440367	0.394495413	0.110091743	0.100917431	0.027522936
lexeme	Chill	Stir	Season	Hull	Pare	Fry	Blanch
	Ice	Marble	Defrost	Crush	Rind	Caramelize	Barbecue
	Congeal	Clot	Bone	Grate	Pod	Coddle	Poach$_1$
	Glaciate	Curdle	Roast	Chop	Scale	Deep-fry	
	Freeze	Tenderize	Husk	Pluck	Sautye	Poach$_2$	
	Cook	Fold	Stew	Dot	Braise	Panfry	
	Work	Beat	Drizzle	Peel	Broil	Frizzle	
	Ferment	Spoon	Filter	Dredge	Scallop	Grill	
	Iodize	Cut	Pound	Grind	Charcoal	Simmer	
	Soak	Deglaze	Bake	Pit	Charbroil	Brown	
	Pickle	Blend	Fillet$_1$	Bolt	Confit	Parboil	
	Marinate		Brush	Gut	Render		
	Salt		Dunk	Toss			
	Whisk		Pipe	Clean			
	Whip			Steam			
				Dice			
				Julienne			
				Mince			
				Slice			
				Bias-slice			
				Panbroil			
				Melt			
				Stone			
				Sift			
				Core			
				Disembowel			
				Eviscerate			
				Baste			

Shred
Sliver
Carve
Butter
Dust
Toast
Devil
Skin
Reduce
Crack
Shuck
Stem
Shell
Fillet_2
Boil

Table 2. Russian lexemes.

i	1	2	3	4	5	6	7
Δ_q	[0.778; 1.799)	[1.799; 2.820)	[2.820; 3.840)	[3.840; 4.862)	[4.862; 5.882)	[5.882; 6.903)	[6.903; 7.925]
m_q	0.154929577	0.23943662	0.197183099	0.183098592	0.154929577	0.056338028	0.014084507
lexeme	охлаждать	мешать	шпиговать	запекать$_2$	варить	калить	пряжить
	замораживать	раскатывать	полосовать	запекать$_1$	ощипывать	бланшировать	
	чистить$_1$	начинять	обваливать	тереть	тушить	коптить	
	ферментировать	прослаивать	размораживать	топить$_1$	жарить	филетировать	
	сбраживать	резать	выжимать	жиловать	томить		
	квасить	мочить	посыпать	солить$_1$	припускать		
	мариновать	замешивать	молоть	шинковать	парить		
	взбивать$_2$	створаживать	глазировать	просеивать			
	мять	засахаривать	засахарить	цедить	вялить		
	солить	крошить	потрошить	пластовать	пластать		
	приправлять	рубить	толочь	перчить	обжаривать		
		фаршировать	лущить	пастеризовать			
		месить	сдобить	засахаривать			
		взбивать$_1$		панировать			
		печь					
		чистить$_2$					
		смазывать					

Appendix II. List of dictionaries

1. *Cambridge Encyclopedia of Language*, edited by David Crystal, 2nd edn. Cambridge: Cambridge University Press, 1997.
2. *Cambridge International Dictionary of English: Guides you to the Meaning*. Cambridge: Cambridge University Press, 1996.
3. *Longman Dictionary of Contemporary English*, director, Della Summers, new edn. Harlow, UK: Longman Dictionaries, 2006.
4. *Longman Dictionary of English Language and Culture*. Harlow, UK: Longman Dictionaries, 1999.
5. *Oxford Advanced Learner's Dictionary of Current English*, A.S. Hornby, 6th edn. Oxford: Oxford University Press, 2000.
6. *Oxford Dictionary and Thesaurus*, edited by Sara Tulloch. Oxford: Oxford University Press, 1993.
7. *Roget's 21st Century Thesaurus in Dictionary Form*, edited by The Princeton Language Institute. Delta Books (Online), 1993.
8. *The New Webster's Comprehensive Dictionary of the English Language. Deluxe edn.* New York: American International Press, 1992.
9. *The Random House Unabridged Dictionary*, 2nd edn, New York: Random House, 1993.
10. *ABBYY Lingvo 12.0*, Machine-readable dictionary. ABBYY Software, 2006.
11. *English–Russian Dictionary of Synonyms*,Y. D. Apresyan. Moskva: Russky Yazyk Publishers, 1979.
12. *Novyi Bolshoi Anglo-Russkii Slovar: Vol. 1–3*, edited by Yu. D. Apresyan. Moskva: Russky Yazyk Publishers, 1999.
13. *Russian Language Dictionary: Vol. 1–4 (Slovar russkogo yazika)*, edited by A. Yvgenyeva, Russian Academy of Sciences, Institute for Linguistic Studies. Moskva: Russky Yazyk Publishers, 1999.
14. *Russkii glagog i ego prichastnye formy. Tolkovo-grammaticheskii slovar*, I. K. Sazonova. Moskva: Russky Yazyk Publishers, 1989.