STUDIES IN LANGUAGE AND MIND 2

SELECTED PAPERS FROM THE 4TH & 5TH WORKSHOP ON *PSYCHOLINGUISTIC, NEUROLINGUISTIC AND CLINICAL LINGUISTIC RESEARCH*

CONTENTS

Dušica Filipović Đurđević
Department of Psychology, Faculty of
Philosophy, University of Novi Sad
dfdurdevic@ff.uns.ac.rs

Aleksandar Kostić
Department of Psychology, Faculty of
Philosophy, University of Belgrade
akostic@f.bg.ac.rs

NUMBER, RELATIVE FREQUENCY, ENTROPY, REDUNDANCY, FAMILIARITY, AND CONCRETENESS OF WORD SENSES: RATINGS FOR 150 SERBIAN POLYSEMOUS NOUNS[*]

Abstract: We collected several measures of ambiguity for 150 Serbian polysemous nouns. Ambiguity measures were derived separately for dictionary senses, and the senses provided by native speakers. In a sense collection task, participants listed all senses of the given word they could think of. Collected senses were categorized in two ways – by preserving fine grained semantic intuition of the speakers as much as possible, and by mapping them onto dictionary categories. In addition, we collected familiarity and concreteness ratings of each dictionary sense, and each sense provided by participants. Based on the senses provided, we calculated the number of senses, the proportion of each sense, entropy and redundancy of sense probability distribution. In order to control for the possible influence of idiosyncratic answers, all ambiguity measures were additionally corrected based on sense frequencies and familiarity ratings. Finally, participants rated word familiarity and word concreteness. The provided measures are to be applied in the research of the processing of polysemous words with a specific accent on the processing effects of meaning uncertainty and balance of sense probabilities. Additionally, they are to help understand the relation between concreteness and polysemy, the relation between semantic intuition and dictionary senses and so forth. All of the collected senses, their frequencies,

---

familiarity and concreteness ratings, as well as lexical ambiguity measures, word familiarity and word concreteness ratings are provided in the supplementary material.

Key words: polysemy, number of senses, entropy, redundancy, sense probability, sense familiarity, sense concreteness

1. Introduction

In the research dedicated to the effect of number of meanings on the processing of polysemous words, number of meanings has been determined in several ways. According to the source they rely on, traditional procedures applied to estimate the number of word meanings could be divided in two groups: the ones that rely on dictionaries and the ones that rely on native speakers. More recently, a third line of procedures has been introduced – the line that encompasses techniques of quantitative linguistics. In that approach, word ambiguity is estimated automatically, based on large language corpora (Landauer and Dumais, 1997; Lund and Burgees, 1997; McDonald, 2000; Schütze, 1998). However, these techniques suffer from numerous problems and have not yet replaced human intuition. Therefore, in this paper, we will focus on traditional procedures (see Filipović Đurđević & Kostić, 2009 for the application of quantitative linguistic approach of distributional semantics to polysemy in Serbian).

1.1 Dictionary meanings

In the first studies dealing with word ambiguity the authors usually relied upon unabridged dictionaries when estimating number of meanings (Gernsbacher, 1984; Jastrzembski and Stanners, 1975; Jastrzembski, 1981; Rodd, Gaskell, & Marslen-Wilson, 2002). This approach has been revived recently with the development of software for meaning collection (Armstrong, Tokowicz, & Plaut, 2012). Dictionary meanings are very suitable for use because they are systematized and accessible. However, there are several problems related to the technique of dictionary based determining of number of

meanings. Firstly, there are large differences among different dictionaries in the way they present word meanings. The origin of these differences lies in the different author's approach to the criterion of grouping of certain word meanings. Consequently, different numbers of meanings are estimated based on different dictionaries. The other problem is the fact that in the dictionaries all the known meanings are listed, among which are often those that are no longer in use, or are known to a very small group of people (e.g. archaisms, localisms, specialized terms, etc.). Gernsbacher (1984) found that even very educated individuals, such as university professors, could not list all the meanings listed in the dictionaries. At the same time, some modern and recent word meanings are not listed in the dictionaries. A study conducted with a purpose to explicitly compare meanings listed in the dictionary with the meanings provided by speakers showed that the two groups differed in both meanings and contents (Lin & Ahrens, 2005). In general, this study showed that a slightly larger number of word meanings were listed in the dictionaries, but only part of dictionary meanings corresponded to the customary meanings. The rest of the meanings listed by the participants were new meanings that have developed in the course of language evolution.

1.2 Meanings provided by native speakers

As an alternative source of data for the number of word meanings, some authors suggested the speakers of the language in question (Azuma & Van Orden, 1997; Borowsky & Masson, 1996; Ferraro & Kellas, 1990; Gawlick-Grendell & Woltz, 1994; Gilhooly & Logie, 1980a; 1980b; Hino & Lupker, 1996; Kellas, Ferraro & Simpson, 1988; Millis & Button, 1989; Rubenstein, Garfield & Millikan, 1970; Twilley, Dixon, Taylor, & Clark, 1994). Millis and Button (1989) proposed the term accessible polysemy to denote the number of meanings familiar to native speakers, while Lin and Ahrens (2005) named it semantic intuition. We could distinguish three groups of techniques that rely on the intuition of native speakers: a) subjective ambiguity rating, b) listing of the first meaning and c) listing of all the familiar meanings.

1.2.1. Subjective ambiguity rating

One way to compare unambiguous and ambiguous words was to ask participants to estimate, on a three point scale, for each word whether it had one, two or three meanings (Borowsky & Masson, 1996; Ferraro & Kellas, 1990; Hino & Lupker, 1996; Kellas, Ferraro, & Simpson, 1988). In addition to providing only the basic information about ambiguity, this procedure had several other disadvantages. For example, Lin and Ahrens (2005) stated that there is a possibility that participants did not think enough about all the meanings when making a decision, at least not to the same extent as in the task with listing of all the meanings. Additionally, the criteria used by the participants when making a decision remain unknown. Finally, in the group of words estimated as ambiguous words, large oscillations in the number of meanings were being neglected.

1.2.2. Listing of the first meaning

Another way to get an approximation of the number of meanings was by asking the participants to list the first meaning they could think of (first meaning metric; Forster & Bednall, 1976; Gawlick-Grendell & Woltz, 1994; Gilhooly & Logie, 1980a; Rubenstein, Garfield & Millikan, 1970; Twilley et al., 1994). Independent evaluators would then analyze participants' responses, classify them and determine the number of different responses. However, the results of the experiments examining the effect of the number of meanings estimated in this way are inconsistent. Depending on the choice of stimuli and the degree of difference in the number of meanings, the ambiguity effect was present in some cases (Rubenstein, Garfield & Millikan, 1970), but not in others (Forster & Bednall, 1976). As Millis and Button (1989) stated, the basic disadvantage of this measure is the fact that by noting down the first meaning the participants can think of, only a small number of dominant meanings is collected, while some less frequent meanings are overlooked. By this procedure, words with one dominant meaning are proclaimed unambiguous words, whereas words with more equally frequent meanings are classified as ambiguous.

1.2.3. Listing of all familiar meanings

In order to solve the problem of neglecting non-dominant meanings present in the procedure of first sense listing, some authors suggested asking the participants to list all the meanings they could think of (Azuma, 1996; Millis & Button, 1989). Afterwards, independent evaluators would classify collected meanings and determine the total number of different meanings listed by all the participants (total meaning metric), and average number of meanings per participant (average meaning metric). In three lexical decision task experiments Millis and Button (1989) tested the three ambiguity measures. Their findings showed that the assessment of number of meanings based on listing of the first meaning was not adequate. Contrary to that, the number of senses assessed on the basis of listing of all the familiar meanings proved as a significant predictor of lexical decision latencies. Significant effect of the number of meanings was observed both for total meaning metric and for average meaning metric. Azuma (1996) suggested the use of total meaning metric, and in support of this idea she stated that it would be impossible for the participant to recollect all of the familiar meanings in a short period of time. She suggested that the set of word meanings should be formed based on all the meanings listed by all the participants, with an additional step in which the participants would rate the familiarity of each of the collected meanings. A study conducted by Azuma showed that familiarity rating of word meanings was a useful supplement to the procedure of full listing of familiar meanings.

Frequencies of individual meanings were often available in the mentioned studies and were used to indicate the existence of the dominant meaning, that is the meaning with the highest frequency. However, Gilhooly and Logie (1980a) suggested a more detailed ambiguity measure based on the overall distribution of frequency of meanings. They named this measure meaning uncertainty (U) and they interpreted it as an average uncertainty of the dominant meaning, which is equivalent to entropy of probability distribution of word meanings. This measure was later adopted by Twilley et al. (1994), and several variations have been proposed in the following years. For

example, Armstrong, Tokowicz, & Plaut (2012) proposed largest relative meaning frequency, or β to describe meaning dominance.

1.3. Current study

When estimating lexical ambiguity the authors of the early studies mostly overlooked the difference between homonymy and polysemy. Starting from the finding that polysemy and homonymy are processed differently (Beretta, Fiorentino, Poeppel, 2005; Klepousniotou, and Baum, 2007; Klepousniotou, Pike, Steinhauer, & Gracco, 2012; Rod, Gaskell, and Marslen-Willson, 2002), we have set our focus exclusively on polysemous nouns. Polysemous nouns are the ones with several related senses (e.g. paper), whereas homonymous nouns have several unrelated meanings (e.g. bank). Unlike homonymous nouns, which take more time to process than unambiguous nouns, polysemous nouns are processed faster.

Having in mind the finding that entropy (MacKay, 2003; Shannon, 1948) has proven as a strong predictor of processing latencies at various levels of descriptions of language (e.g. Baayen, Feldman, and Schreuder, 2006; Baayen, Milin, Filipovic Đurđevic, Hendrix, and Marelli, 2011; Balling, and Baayen, 2012; Milin, Filipović Đurđević, and Moscoso del Prado Martìn, 2009; Mosocoso del Prado Martin, Kostić and Baayen, 2004; Tabak, Schreuder, and Baayen, 2005; Wurm, Ernestus, Schreuder, and Baayen, 2006), we have set as the basic goal of this paper the estimation of the entropy of word sense probability distribution (1).

(1)

$$H(w) = -\sum_{i=0}^{n} p_i \cdot \log p_i \quad (1)$$

In (1) H denotes entropy of the polysemous word w, index i stands for different senses of word w, p_i denotes the proportion (relative frequency) of the given sense of w, and n denotes the number of senses of w. This measure provides a more detailed index of word

ambiguity (or degree of ambiguity [U] as suggested by Gilhooly and Logie [1980a]). When compared to the number of senses that has been traditionally applied in polysemy research, the added information that is included in entropy concerns the balance of sense probabilities. Entropy of sense probability distribution can be interpreted as uncertainty of senses. It is influenced by the number of senses in such a way that a larger number of senses leads to larger entropy, that is a larger degree of uncertainty of the true sense of the word (with logN being the theoretical maximum). However, it is also influenced by the balance of sense probabilities in that words with balanced probabilities of senses carry greater sense uncertainty, that is larger entropy. Words with a dominant sense, that is unbalanced sense frequencies carry less uncertainty of the true sense of the word. This added information can be described independently of the number of senses via the Information Theory measure of redundancy (2).

(2)

$$T(w) = 1 - \frac{H(w)}{\log N} \quad (2)$$

In (2) T(w) stands for the redundancy of the polysemous word w, H(w) stands for its entropy, and N denotes the number of senses of word w. The larger the redundancy, the less balanced the distribution of sense probabilities, that is the less the uncertainty of the true sense of the word.

The approach of describing polysemy as sense uncertainty brings an advantage, as suggested by Gilhooly and Logie (1980a), as it offers a more detailed description of the degree of ambiguity. Additionally, separate quantification of the two sources of sense uncertainty, namely number of senses and redundancy (balance of sense probabilities), brings the additional advantage of separate investigation of the influence that these can have on the processing of polysemous words. Polysemous words need not be categorized into words with a dominant sense and words with a balanced sense, as was the case previously (Duffy, Morris, & Rayner, 1988; Klepousniotou, Titone, Romero,

2008; Simpson, 1994; Swinney, 1979; Whitney, Jefferies, & Kircher, 2011) – the degree of balance of sense probabilities can be controlled or investigated in a more detailed manner.

Our approach is similar to the one of Gilhooly and Logie (1980a). However, unlike their study in which the participants listed the first meaning they could think of, in this study we asked participants to list all the senses they could think of (as suggested by Azuma, 1996) and offered several corrections of the estimated number of senses based on several criteria. Also, in addition to calculating entropy of the dictionary senses listed by the participants, as was done by Gilhooly and Logie (1980a), we calculated entropy of the raw senses listed by the participants (without categorizing them based on the dictionary senses). Finally, in addition to calculating entropy, we calculated redundancy of the distribution of sense probabilities.

With all of the previously described approaches in mind, the estimation of the number of senses of Serbian polysemous words was performed in several ways. Firstly, based on the senses stated in an extensive dictionary of Serbian a sample of 150 polysemous Serbian nouns was excerpted. For each noun we collected familiarity ratings (subjective frequency) and word concreteness ratings. After that, for each of these nouns, we collected all the senses that the participants, native speakers of Serbian language, could think of. Additionally, the collected senses were categorized in compliance with the senses listed in the dictionary. The number of senses was estimated in two ways. On the one hand, the number of senses listed by the participants was determined, and on the other hand, the number of dictionary senses that were being listed by the participants was determined. In addition to the number of senses (N), frequencies of listing each sense were determined. Based on the number of participants listing a specific sense (sense frequency), we derived a proportion (relative frequency) of the sense in question in relation to other word senses (p). Entropy and redundancy are derived from determined proportions by applying (1) and (2).

In the next step, we collected familiarity ratings separately for the senses listed in the dictionary and the senses listed by the participants. Based on these ratings, the estimated number of senses

was corrected by excluding from the set of senses those senses that were not familiar to the majority of the participants. Additionally, we collected concreteness ratings for individual senses (separately for the ones listed in the dictionary and the ones listed by participants).

A summary of the collected ratings can be found in the Appendix and the supplementary material containing the full dataset can be accessed online.[1]

2. Dictionary based number of senses

In this study, the dictionary based number of senses was used as a starting point. This measure has been used in a large number of studies, in spite of the numerous downsides that are related to this technique of estimating the number of senses (Armstrong, Tokowicz, & Plaut, 2012; Gernsbacher, 1984; Jastrzembski, 1981; Jastrzembski and Stanners, 1975; Lin and Ahrens, 2005; Rodd et al., 2002).

2.1 Method

2.1.1. Stimuli

Based on *Rečnik Matice srpske* dictionary (the most extensive completed dictionary of Serbian), 150 Serbian nouns were selected that have several senses listed in one dictionary entry, that is which satisfy the criterion of polysemy stated by linguists (it is common to consider separate entries as separate lemmas [Rodd et al., 2002]). Only the words that do not overlap with various inflected forms of other

---

[1] All collected senses and their associated measures can be found following this link: https://drive.google.com/file/d/0B0HHGsBbpIrleVY1U3l0LTRWMTA/view?usp=sharing

Per-word summary of collected measures can be found following this link: https://drive.google.com/file/d/0B0HHGsBbpIrlN29ja0cxQzAycWM/view?usp=sharing

word classes were selected to insure the investigation of strictly polysemous nouns (for example, the noun *baza* [base] is excluded due to being a homograph with one of the present tense forms of the verb *bazati – on baza* [to wonder around – he wonders around]). Therefore, all of the selected stimuli were strictly polysemous Serbian nouns. The words were selected to span as high as possible a range of number of senses, and lemma frequencies obtained from a frequency dictionary (Kostić, 1999). An attempt was made to decorelate number of senses and lemma frequencies by matching the words for their lemma frequencies (as closely as possible) accross the categories of words with a given number of senses

2.1.2. Procedure

The number of senses was established in two ways. Firstly, by counting all of the senses, regardless of the grouping assigned by the authors of the dictionary. For example, this way, the word *gluma* (an act) had four senses. Also, the number of senses was determined by counting the clusters of senses, as stated in the dictionary. This way, the same word had three senses (Table 1).

Table 1. Rečnik Matice srpske dictionary description of word *gluma* (an act).

gluma (an act)	1	skill of an actor (*His act was very good in that movie.*)
	2a	a theatre play (*I went to the theatre to see the new act.*)
	2b	theatre art (*The act in Belgrade is very good.*)
	3	(figurative) pretending (*Do not trust him, it is all an act!*)

2.2. Results and discussion

Distributions of the two measures of number of senses are shown in Table 2. The selected polysemous nouns had on average 6.71 senses when all of the senses were taken into account, and 4.51 clusters of senses. There was a significant correlation between the two counts (r=.748, p<.001).

Table 2. The distribution of the number of senses listed in the dictionary. Rows refer to number of senses given in the first column; the second column illustrates how many words have the given number of individual senses and the third column contains average lemma frequencies of those words; the fourth column contains number of words that have a given number of sense clusters and the fifth column contains their average lemma frequencies. The two final rows contain the mean and the standard deviation of values presented in the respective columns.

	Individual senses		Clusters of senses	
Number of senses	Number of words	Mean lemma frequency	Number of words	Mean lemma frequency
1			2	79.00
2	12	241.17	21	187.67
3	14	127.71	33	140.33
4	20	143.15	37	217.68
5	16	197.81	15	226.20
6	21	202.29	16	190.56
7	16	163.06	8	230.88
8	9	200.22	9	344.00
9	12	271.75	7	271.00
10	6	226.83		
11	10	293.50	1	176.00
12	4	351.50	1	365.00
13	4	287.00		
14	3	198.67		
15	1	60.00		
18	2	235.50		
M	6.71		4.51	
SD	3.42		2.14	

Levene's test of homogeneity of variance demonstrated that frequency was equally distributed across the categories of number of senses. However, in spite of this, there was a significant correlation between lemma frequency and the number of all of the senses listed in the dictionary (r=.184, p<.05).

3. Collecting word familiarity ratings and ratings of the senses listed by native speakers

Having in mind all the downsides of using a dictionary in estimating the number of senses, we also collected measures of so called accessible polysemy (Millis & Button, 1989), that is we conducted a study aimed at estimating the number of senses that are accessible to participants. Having in mind the downsides and the unreliability of the technique in which participants list only the first meaning they can think of, in this study participants were asked to list all of the senses they could think of (Azuma, 1996; Millis & Button, 1989).

3.1 Method

3.1.1. Participants

Seventy-four first year and second year students from the Department of Psychology, Faculty of Philosophy at the University of Belgrade participated in the study. Some of them also took part in the remaining studies that we reported in this manuscript.

3.1.2. Stimuli

The words selected from *Rečnik Matice srpske* dictionary were divided in five groups, in such a way that the average number of dictionary senses was equal across groups. Word groups were printed in separate ten-page booklets. In each booklet, words were printed in five random orders. Next to each word, a seven point word familiarity

rating scale was printed, and an empty space for listing senses was placed.

3.1.3. Procedure

Each participant filled only one booklet, that is listed the senses for one group of 30 words. The first page of the booklet contained a detailed instruction and an example. The task of the participant was to read the word and rate its familiarity (how often he/she encountered the word) on a seven-point scale. Number one marked a word that was completely unfamiliar, while number seven marked a word that was very familiar:

COMPLETELY UNFAMILIAR WORD   1   2   3   4   5   6   7   VERY FAMILIAR WORD

After that, the task was to list all of the senses of a word the participant could think of, using the empty space below each word (various concepts denoted by a word, various usages of the concept, etc). Participants were advised to rely on as many means as possible while listing the senses, such as the definition of meaning, a synonym, a sentence illustrating the usage, and so forth.

3.2. Results and discussion

All of the words were rated as highly familiar. The average familiarity for 150 polysemous nouns was 6.28 units of the seven point scale (SD=.48). Senses of each word were collected based on the descriptions of 17 to 19 participants (group 1: N=17; group 2: N=18; group 3: N=19; group 4: N=18; group 5: N=19). We derived measures of the total number of senses listed by the participants and the average number of senses per participant. Distributions of the collected measures are listed in Table 3. Total and average number of senses was determined in two ways. On the one hand, we calculated the number of raw, uncategorized senses listed by the participants, and on the other

hand, we calculated the number of senses that matched one of the senses listed in the dictionary.

The raw number of senses was determined by considering each of the senses listed by the participants as a separate sense (Azuma, 1996). Senses were kept separate even when they were a more specific instance of a more general sense. This decision was made in order to preserve a fine grained semantic distinction. For example, very often there were notable differences between the characteristics of the objects denoted by a general, and those denoted by a more specific instance of a particular word sense. In accordance with this view is Azuma's finding that a large number of participants stated separately general senses and their specific instances. This rule was broken only when it was obvious that the participant defined the same sense in several manners. In order to avoid the possibility of listing idiosyncratic senses, we calculated the number of senses listed by more than 10% of the participants. In our case, this means that senses listed by only one participant were excluded from the list.

In addition to the number of raw senses listed by the participants, we calculated the number of dictionary senses listed by the participants. This was done by categorizing the raw senses listed by the participants according to the dictionary, that is by matching each sense listed by the participants with an adequate dictionary entry. After that, we counted the dictionary entries that appeared in the participants' answers.

Participants listed 2.94 senses on average, 2.23 of which were listed in the dictionary. On the one hand, a slightly larger number of the raw senses listed by the participants was a consequence of the applied principle of keeping the variety of answers in counting the word senses. This principle led to a more fine grained, or higher "resolution", and consequently to a larger number of senses. On the other hand, this difference was a consequence of the existence of senses listed by the participants but not appearing in the dictionary. The two estimated measures of number of senses were moderately correlated (r= .49, p<.01).

The average number of senses listed per word was 12.51. After eliminating senses listed by less than 10% of participants (i.e. listed by only one participant) the average total number of senses listed by

participants was 7.97. The correlation of the number of senses listed by the participants before and after eliminating idiosyncratic answers was positive, and statistically significant (r=.80, p<.01). We recorded a significant correlation between the total and the average number of senses per participant (before eliminating rare answers: r=.76, p<.01; after eliminating rare answers: r=.78, p<.01). The number of senses listed per participant was moderately correlated with the number of senses listed in the dictionary (before eliminating rare answers: r=.40, p<.01; after eliminating rare answers: r=.33, p<.01).

After categorizing the senses according to the dictionary entries, the average number of listed senses was 4.41. After eliminating the senses listed by less than 10% of the participants, the average number of listed senses decreased to 3.97. The two measures of number of senses were highly correlated (r=.95, p<.01). On average, participants listed 2.23 senses listed in the dictionary. The total number of dictionary senses listed by the participants and the average number of dictionary senses per participant were moderately correlated (r=.56, p<.01, regardless of eliminating rare answers). As expected, the correlation between the number of senses listed in the dictionary and the number of dictionary senses listed by participants was higher than the correlation between the number of senses listed in the dictionary and the number of raw senses listed by the participants (before eliminating rare answers: r=.82, p<.01; after eliminating rare answers: r=.77, p<.01).

Table 3. The distribution of the number of senses listed by the participants, prior to categorization (left hand side) and after being categorized according to the dictionary (right hand side). Rows mark the number of senses listed in the first column, and cells contain the number of words with a given number of senses: the total number of listed senses (Total), the number of senses listed by more than 10% of the participants (Total>10%), the average number of senses per participant (M), the most frequent number of senses per participant (mode), and the median number of senses per participant (median). The final two rows contain the mean and the standard deviation of values listed in the respective columns. For example, number 59 in the second row and the fifth column states that there were 59 words for which the mode of the number of senses listed by the participants was 2 (i.e. that there were 59 words for which the participants most frequently listed 2 senses).

| Number of senses | Number of words with the given number of senses (as listed in the first row) | | | | | | | | | |
| | Raw senses listed by participants | | | | | Dictionary senses listed by participants | | | | |
	Total	Total >10%	M	Mode	Median	Total	Total >10%	M	Mode	Median
1			1	1	1	2	2	14	15	14
2			37	59	45	25	30	94	95	95
3		4	83	61	77	32	33	37	34	35
4	1	8	25	25	24	31	37	5	5	6
5	2	16	4	3	3	23	26		1	
6	11	18		1		12	8			
7	7	31				11	7			
8	14	22				5	5			
9	10	13				6	1			
10	12	10				2	1			
11	15	9				1				
12	13	7								
13	13	3								
14	9	6								
15	9	1								
16	8	1								
17	4	1								
18	3									
19	4									
20	5									
21	1									
22	1									
23	1									
24	1									
25	5									
34	1									
M	12.51	7.97	2.94	2.82	2.85	4.41	3.97	2.23	2.21	2.20
SD	5.07	2.84	.70	.84	.73	2.08	1.70	.56	.69	.65

In addition to counting the number of senses, the applied procedure of collecting senses enabled us to calculate the frequency of each sense, i.e. the number of participants who listed a given sense. Based on frequency, we calculated the proportion of each sense, relative to all listed senses. In the next step, based on the proportions, we calculated Information Theory measures describing the characteristics of the whole distribution. These measures were entropy and redundancy. Considering the fact that we counted the number of senses in several ways, both entropy and redundancy were calculated for each of the obtained number of senses, that is for: a) raw senses listed by the participants, b) raw senses listed by more than 10% of the participants, c) dictionary senses listed by the participants, and d) dictionary senses listed by more than 10% of the participants.

The number of senses listed in the dictionary and the number of senses listed by the participants, along with corresponding entropies and redundancies of sense probability distributions are listed in the supplementary data.

4. Collecting familiarity ratings for dictionary senses

In spite of the numerous downsides, the number of senses listed in the dictionaries should not be discarded. The classification criteria applied in the dictionaries reflect important aspects of linguistic semantic theories. Taking into account the significance of the theoretical basis for estimating the number of senses, we conducted a study aimed at overcoming some of the downsides of the dictionary based estimation of the number of senses. The most common critique refers to the fact that dictionaries list many of the senses that are unfamiliar to average speakers (Gernsbacher, 1984; Lin & Ahrens, 2005). Overcoming this downside by categorizing the senses listed by the participants according to dictionary senses requires a high level of linguistic competence or expertise and introduces new problems. Therefore, the estimated numbers of senses were corrected by collecting sense familiarity ratings. We conducted a study in which participants rated the familiarity of each sense listed in the dictionary. This way, the number of senses listed in

the dictionary was transformed to the number of dictionary senses that are familiar to the majority of the participants.

4.1. Method

4.1.1. Participants

Ninety-one first year students from the Department of Psychology, Faculty of Philosophy at the University of Belgrade participated in the study. The participants from this study partially overlapped with participants from other studies reported in this paper.

4.1.2. Stimuli

One hundred and fifty words selected in the first phase of the study were divided into four groups in such a way as to keep the average number of senses (as listed in the dictionary) equal across the four groups. The words and senses were printed in three random orders in separate booklets, making nine random orders in total.

4.1.3. Procedure

Each participant filled one booklet, i.e. rated one 35-word group. Each page of the booklet consisted of three columns. The first column contained a word, the second column contained the descriptions of each of the senses taken from the dictionary *Rečnik Matice srpske* (one description per row), while the third column contained a seven point scale printed next to the sense description. The first page of the booklet contained a detailed instruction and an example. The task of the participant was to read all of the listed senses and use the seven-point scale to rate the familiarity of a given sense (how often they have encountered it). If a word sense was very familiar, that is if they have encountered a given word in a given sense often, a 7 was to be circled. If the sense was partially familiar, that is, a given word was sometimes encountered in a given sense, a 3, or 4 was to be circled. On the other hand, if a given sense of a word was completely unfamiliar, that is, if

they have never encountered a given word in a given sense, a 1 was to be circled. The participants were advised to use the whole range of the scale:

COMPLETELY UNFAMILIAR WORD SENSE   1 2 3 4 5 6 7   VERY FAMILIAR WORD SENSE

4.2. Results and discussion

Sense familiarity measures of each word were derived based on the ratings of 20 to 27 participants (group 1: N=27; group 2: N=21; group 3: N=23; group 4: N=22). The distributions of the collected measures are listed in Table 4.

The average number of the senses that are familiar to the participants was calculated by determining the number of senses that were rated above 1 on the familiarity scale. This was done for each participant separately, and after that three measures of central tendency were derived for the number of familiar senses (average, mode, and median). The average number of senses that are familiar to the participants was 5.82, which in comparison with the average number of senses listed by the participants (4.41) was in accordance with the assumption that the participants were not listing all of the senses they are familiar with (cf. Azuma, 1996). In spite of that, the two measures were positively correlated (r=.65, p<.01).

Table 4. The distribution of the total number of senses, and the average number of senses per participant (prior to categorization: left hand side; after being categorized according to dictionary senses: right hand side) based on the sense familiarity judgment, obtained by applying three criteria: counting the senses with the mean sense familiarity rating greater than or equal to 2 (M>=2), counting the senses with the most frequent sense familiarity rating larger than 1 (mode>1), and counting the senses rated by more than 50% of the participants by above 1 sense familiarity (median>1). The rows represent the number of senses listed in the first column, and the cells contain the number of words that have the given number of senses. The final two rows contain the mean and standard deviation of values listed in respective columns.

	Number of words											
	Uncategorized senses listed by participants						Dictionary senses					
	Total number of senses			Average number of senses per participant			Total number of senses			Average number of senses per participant		
Number of senses	M ≥2	Mode >1	Median >1	M ≥2	Mode >1	Median >1	M ≥2	Mode >1	Median >1	M ≥2	Mode >1	Median >1
1				3	17	12						
2				4	2	1	14	18	14	14	14	14
3				2			17	22	18	16	18	16
4	1	1	1	2	1	1	21	32	21	26	22	24
5	4	11	4	8	3	4	20	20	24	25	24	23
6	11	11	12	18	12	15	23	15	19	20	21	24
7	5	7	6	10	7	6	12	7	11	10	8	8
8	16	15	16	18	16	18	6	11	7	8	8	6
9	11	14	10	14	10	12	11	9	13	12	13	14
10	15	16	14	4			8	6	7	6	6	8
11	13	12	15	8	14	13	6	3	4	6	7	4
12	12	5	11	7	8	9	6	6	5	4	3	5
13	11	11	10	9	12	10	3		4	2	5	3
14	10	15	11	13	12	15	2		2			
15	10	4	10	8	8	7						
16	7	8	6	4	6	6						
17	5	5	6	6	5	6		1		1	1	1
18	4	3	3	3	4	3	1		1			
19	2	2	2	1	3	3						
20	3	1	3	1								
21	1	1	1		2	1						
22	1	1	1			1						
23	1	3	2	4	2	1						
24	2	1	2	2	2	3						
25	4	2	3		3	2						
30				1								
33						1						
34		1										
35	1		1		1							
M	12.25	11.53	12.16	10.45	10.86	10.96	6.23	5.47	6.15	5.82	6.05	5.98
SD	5.05	5.00	5.04	5.26	6.14	5.73	3.19	2.87	3.18	2.89	3.06	2.96

For each dictionary sense, we derived three measures of central tendency: average, median, and mode of participant's familiarity

judgments. After that, based on each of these measures of central tendencies, we derived a new, corrected number of senses. In case of average-based measures, we counted only senses with a mean familiarity rating equal to or above 2.00. In the case of mode and median-based measures, we counted only the senses with a mode or median familiarity judgment above 1. In other words, we counted only the senses which were rated above 1 by at least half of the participants (median), and only senses that were not rated with 1 in the majority of cases (mode).

As expected, the average number of dictionary senses that was familiar to the majority of the participants was less than the number of senses listed in the dictionary. The resemblance of the two measures was highest in case of counting the senses based on the criterion of average ratings (6.23). The resemblance was weaker in the case of median (6.15), and was the weakest in the case of mode (5.47). Mode was the most strict criterion in accepting the senses familiar to participants.

The comparisons of the three derived measures of the number of dictionary senses familiar to the participants (based on the three measures of central tendency) revealed high positive correlation coefficients between each of the pairs (r>.95, p<.01). In addition to that, a high positive correlation was observed between the number of senses listed in the dictionary, and the number of dictionary senses with an average familiarity rating of 2 or higher (r=.96, p<.01), as well as with the number of dictionary senses with a median familiarity rating above 1 (r=.96, p<.01). The correlation coefficient was somewhat lower in the case of the number of senses with a mode familiarity rating above 1 (r=.85, p<.01). The three measures were also correlated with the number of dictionary senses listed by the participants. In this case, when compared to the correlation coefficients with the number of senses listed in the dictionary, the correlation coefficient was slightly lower for the senses selected by applying the average-based criterion (r=.86, p<.01), and the median-based criterion (r=.86, p<.01), and remained almost unchanged in the case of the mode-based criterion (r=.88, p<.01).

The correlation coefficient between the average familiarity of word senses and the familiarity rating of the corresponding word was not significant. Word familiarity judgments were correlated only with familiarity judgments of the dominant sense (r=.32, p<.01).

The average familiarity judgments of the dictionary senses were positively correlated with the dictionary sense frequencies, that is, the number of participants who listed a given dictionary sense (r=.68, p<.01). In general, familiar senses were more frequently listed (Figure 1). However, there was a large number of low frequency senses that were rated as highly familiar. We could assume that the observed correlation would increase if the number of participants were increased.

Figure 1: The relation of the relative sense frequencies and the average sense familiarity judgments obtained for the senses listed in the dictionary

5. Collecting familiarity ratings for the senses listed by native speakers

Given that while collecting the raw sense listed by the participants an effort was made to preserve the semantic intuition of the participants as closely as possible, there was a risk of the presence of idiosyncratic senses in the collected sample. To ensure that all of the senses on which the analyses were to be based were familiar to the majority of

participants, we conducted a study in which the participants rated the familiarity of each of the senses previously collected.

5.1. Method

5.1.1. Participants

Eighty-five first year students from the Department of Psychology, Faculty of Philosophy at the University of Belgrade participated in the study. Some of the participants who took part in this study also participated in other studies reported here.

5.1.2. Stimuli

One hundred and fifty words selected in the first phase of the study were divided in four groups, in such a way to keep the average number of senses equal across the four groups. The words and senses were printed in three random orders in separate booklets, making nine random orders in total.

5.1.3. Procedure

Each participant filled one booklet, that is, rated one 35-word group. The task was the same as in rating the familiarity of the senses listed in the dictionary. The only difference was that, instead of the dictionary senses, the raw senses listed by the participants in the first stage of the study were printed next to each word. Along with these senses, five absurd senses were included in the list, in order to control for the potential random filling of the test by the participants.

5.2. Results and discussion

Seven participants who rated either one of the control, nonsense senses as familiar (above 1 on the familiarity scale) were excluded from the sample. Sense familiarity measures of each word were derived based on the judgments of 17 to 22 participants (group 1: N=19; group 2:

N=17; group 3: N=20; group 4: N=22). The distributions of collected measures are listed in Table 4.

The measures of the number of senses listed by the participants, which were familiar to the majority of participants, were derived in the same way as the measures of the number of familiar dictionary senses (previous section).

The average number of raw senses listed by the participants that were familiar to the participants was 10.45, which was higher than the average number of raw senses listed by the participants. However, these two measures were highly correlated (r=.88, p<.01).

The average total number of raw senses that the participants were familiar with was only slightly below the average total number of the raw senses listed by the participants. This number was lower only if the mode-based criterion was applied in the selection of the familiar senses.

High correlation coefficients were obtained for each of the pairs of the three derived measures of the number of familiar raw senses listed by the participants (r>.98, p<.01). In addition to that, we obtained a high correlation coefficient between the raw number of senses listed by the participants and the number of raw senses listed by the participants with a mean sense familiarity of 2 or higher (r=.99, p<.01), as well as the number of senses listed by the participants with a median sense familiarity above 1 (r=.99, p<.01). The correlation coefficient was slightly lower, but also still very high in the case of the number of senses with a mode sense familiarity above 1 (r=.97, p<.01). The three measures of the number of familiar senses were also correlated with the number of raw senses listed by more than 10% of the participants. In this case, the correlation coefficient was the same for all three measures and slightly lower (r=.80, p<.01). A high correlation between the number of listed senses and the number of familiar senses, as well as a decrease in correlation in the case of the number of senses listed by more than 10% of the participants pointed to the absence of idiosyncratic senses in the collected sample. The remaining downside of counting the raw senses listed by the participants was the possibility that a sense listed by only one participant was not distinct enough to be treated as a separate sense.

Word familiarity judgments were positively correlated with the mean sense familiarity judgments (r=.25, p<.01). Although the

correlation coefficient was significant for mean sense familiarity ratings of the subordinate senses (r=.22, p<.01), it was higher in case of the dominant sense familiarity judgments (r=.30, p<.01).

The average sense familiarity judgments were positively correlated with sense frequency, that is the number of participants listing a sense (r=.68, p<.01). In general, the senses with higher familiarity ratings were listed by a larger number of participants (Figure 2). However, there was a large number of low frequency senses that were rated as highly familiar. This was probably due to the great variety of answers produced by the participants.

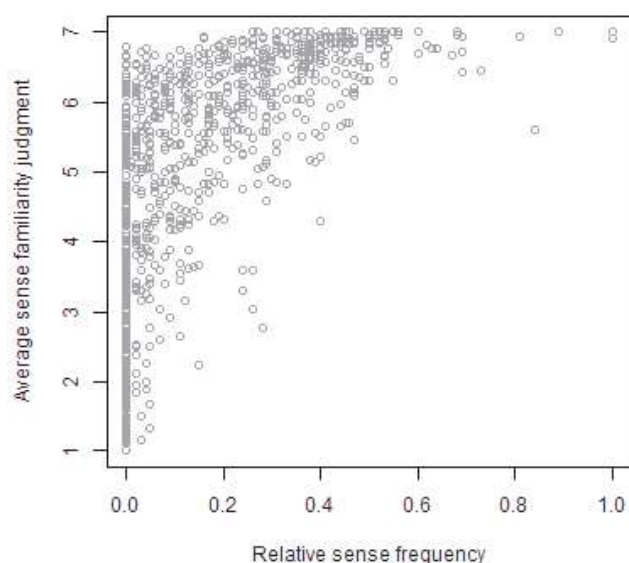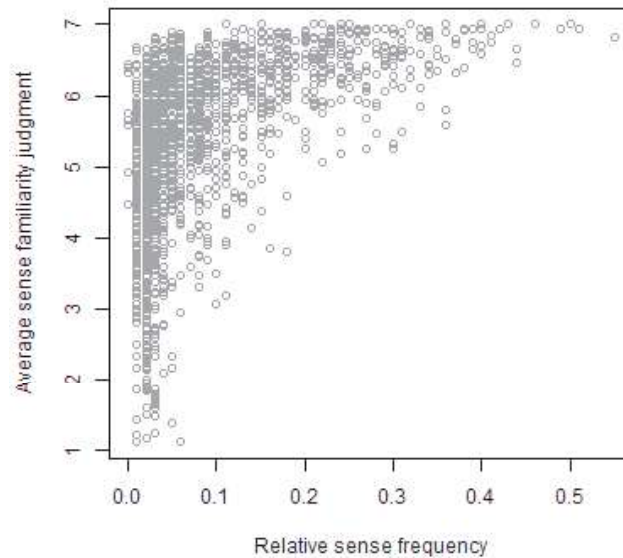Figure 2: The relation of the relative sense frequencies and the average sense familiarity judgments obtained for the senses listed by the participants

6. Collecting word concreteness ratings

It has been demonstrated that concreteness of word meaning affected processing time. Words denoting objects or beings that could be experienced through the senses (seeing, hearing, touching, etc.) are recognized faster than words denoting abstract concepts (Paivio, 1986;

Schwanenflugel, 1991). Therefore, in order to control for the effect of concreteness, we assessed the word concreteness ratings.

6.1. Method

6.1.1. Participants

Forty-three first year students from the Department of Psychology, Faculty of Philosophy at the University of Belgrade participated in the study. These participants also took part in some of the remaining studies reported here.

6.1.2. Stimuli

One hundred and fifty words selected in the first phase of the study were divided into two groups, and printed in three random orders in separate booklets.

6.1.3. Procedure

Each participant filled only one booklet, i.e. rated one group of 75 words. The first page of the booklet contained detailed instructions and an example. The instructions were formulated based on Paivio, Yuille, and Madigan (1968) and their definition of abstractness as the absence of sensory experience. The task of the participants was to read the word and rate the extent of the possibility to experience the object denoted by a word using the senses, that is, to rate its concreteness. They were required to do so by circling the right value on the scale ranging from 1 to 7. Number 1 referred to a very abstract concept, something one can not see, hear, smell, or touch (e.g. 'truth'), while number 7 referred to something very concrete, something one can see, hear, smell, touch, etc. (e.g. 'pig'). Participants were advised to use the whole range of the printed scale:

ABSTRACT 1 2 3 4 5 6 7   CONCRETE

Although imageability is frequently assessed along with concreteness, in this study we opted only for concreteness. We decided to do so by taking into consideration the finding that the correlation between concreteness and imageability is typically high (e.g. r=.83 in Paivio et al., 1968). Additionally, it has been shown that participants typically rate imageability by relying only on visual modality (Connell & Lynott, 2010), whereas all sensory modalities contribute to the representation of word meaning (Filipović Đurđević, Popović Stijačić, & Karapandžić, 2016; Lynott & Connell, 2009; 2010; 2013).

6.2. Results and discussion

For each word, concreteness judgments were estimated based on the answers of around 20 participants (group 1: N=23; group 2: N=21). Average word concreteness of the tested nouns was 4.93 units of the seven-point scale (SD=1.54). More than two thirds of the tested words were rated as concrete, while less than a third was rated as abstract. Word concreteness judgments correlated significantly only with word familiarity ratings (r=.24, p<.01) and the number of dictionary senses listed by the participants (r=-.23, p<.01). However, a significant correlation coefficient was obtained for standard deviation of word concreteness judgments and all of the estimated measures of number of senses (e.g., in the case of the number of raw senses listed by the participants: r=.19, p<.05). This finding matched our expectation that inconsistencies in concreteness judgments would increase with an increase in number of senses based on which one can rate concreteness. In accordance with this interpretation, the participants were informing us about the problems they had while deciding which sense they should rate while rating the concreteness of an ambiguous word.

7. Collecting concreteness ratings for dictionary senses

In addition to the study in which word concreteness ratings were collected, we conducted a study in which participants rated the concreteness of individual senses listed in the dictionary.

7.1. Method

7.1.1. Participants

Eighty-two first year students from the Department of Psychology, Faculty of Philosophy at the University of Belgrade participated in the study. As with the previous studies, this group of participants partially overlapped with participants from the remaining studies.

7.1.2. Stimuli

One hundred and fifty words selected in the first phase of the study were divided into four groups in such way that the average number of dictionary senses was equal across groups. Each group of words was printed in three random orders in separate booklets. At the same time, word senses were printed in three random orders, making nine random orders in total.

7.1.3. Procedure

Each participant filled one booklet, i.e. rated one 35-word group. Each page of the booklet consisted of three columns. The first column contained a word, the second column contained the descriptions of each of the senses taken from the dictionary *Rečnik Matice srpske* (one description per row), while the third column contained a seven point scale printed next to the sense description. The first page of the booklet contained detailed instructions and an example. The task of the participants was to read the word and rate the extent of the possibility to experience the object denoted by the word sense using the senses

(sight, smell, touch…), i.e. to rate its concreteness. They were expected to do this by circling the right value on the scale ranging from 1 to 7. Number 1 referred to a very abstract concept, something one could not see, hear, smell, or touch (e.g. 'truth'), while number 7 referred to something very concrete, something one could see, hear, smell, touch, etc. (e.g. 'pig'). Participants were advised to use the whole range of the printed scale:

ABSTRACT WORD SENSE   1 2 3 4 5 6 7   CONCRETE WORD SENSE

7.2. Results and discussion

For each word, sense concreteness judgments were estimated based on the ratings of around 20 to 21 participants (group 1: N=21; group 2: N=20; group 3: N=20; group 4: N=21). We tested the reliability of the obtained ratings by splitting participants into two groups and looking at the correlation between the average sense concreteness ratings obtained in them. Our results revealed a high positive correlation between the two groups (r=.89, p<.01), as well as a high positive correlation between each of the groups and the global averages (r=.97, p<.01). This provided us with the information that the collected judgments were stable across participants.

Word concreteness ratings and average sense concreteness ratings were positively correlated: r=.68, p<.01. However, this correlation was a concequence of the high correlation between the word concreteness ratings and the dominant sense ratings: r=.71, p<.01. No significant correlation was recorded between word concreteness ratings and the average concreteness ratings of the subordinate senses. This finding indicated that during the process of rating word concreteness, participants were mostly relying on the dominant sense.

We recorded a significant positive correlation between sense concreteness ratings and sense frequencies, i.e. the number of participants who listed a sense in the first phase of the study: r=.36, p<.01. A similar relation was recorded in the case of sense concreteness ratings and sense familiarity ratings: r=.28, p<.01. The

participants were more frequently listing concrete senses. At the same time, they demonstrated a tendency to rate concrete senses as more familiar.

8. Collecting concreteness ratings for senses listed by native speakers

In addition to the study in which participants rated the concreteness of the senses listed in the dictionary, we conducted a study in which participants rated the concreteness of the senses listed by native speakers, which were collected in the first phase of the study.

8.1. Method

8.1.1. Participants

Sixty-five first year students from the Department of Psychology, Faculty of Philosophy at the University of Belgrade participated in the study. These participants also took part in some of the other studies we reported in this paper.

8.1.2. Stimuli

One hundred and fifty words selected in the first phase of the study were divided into four groups, in such way that the average number of dictionary senses was equal across groups. Each group of words was printed in three random orders in separate booklets. At the same time, word senses were printed in three random orders, making nine random orders in total.

8.1.3. Procedure

Each participant filled one booklet, that is rated one 35-word group. The task was identical to the one described in the previous

section, the only difference being that instead of dictionary descriptions of the senses, the descriptions listed by the participants were printed.

8.2. Results and discussion

For each word, sense concreteness judgments were estimated based on the ratings of around 15 to 17 participants (group 1: N=17; group 2: N=16; group 3: N=15; group 4: N=17). As in the previous section, we split participants into two groups and observed a high positive correlation between the averages obtained in the two groups (r=.87, p<.01), as well as between averages from each of the groups and global averages (r=.97, p<.01; r=.96, p<.01).

There was a significant correlation between word concreteness ratings and average sense concreteness ratings: r=.68, p<.01. A slightly higher correlation coefficient was recorded between word concreteness ratings and dominant sense concreteness ratings: r=.73, p<.01. The correlation coefficient between word concreteness ratings and average concreteness ratings of the subordinate senses was lower: r=.17, p<.05. As in the case of the dictionary senses, we could infer that the participants made the word concreteness judgments based on the dominant sense of a word.

A moderate positive correlation was obtained between sense concreteness ratings and sense frequencies, that is, the number of participants listing a sense: r=.30, p<.01. A slightly lower, but significant correlation coefficient was obtained in the case of sense concreteness ratings and sense familiarity ratings: r=.21, p<.01. Based on this, we can conclude that concrete senses were listed more frequently and rated as more familiar by participants.

9. General discussion

We conducted a series of studies aiming at collecting several ambiguity measures. Based on the senses listed in the dictionary *Rečnik Matice srpske*, we selected 150 polysemous Serbian nouns. The selected words were presented in several surveys. Firstly, we collected

all of the senses that the participants, native speakers of Serbian could think of. Based on the collected sample, we formed two lists of senses, which were subjected to further research. On the one hand, we formed a list of raw, uncategorized senses listed by the participants, and on the other hand, by categorizing the raw senses according to the dictionary, we formed a list of dictionary senses listed by the participants. For each of the two lists, we determined the total number of senses, and the average number of senses per participant. In order to control for the potential influence of idiosyncratic senses, the number of senses was corrected by excluding all of the senses listed by less than 10% of participants, that is, listed by only one participant. In addition to that, for each of the lists of senses, we collected familiarity judgments of individual senses. After that, the number of senses was alternatively corrected by excluding the senses that were unfamiliar to the majority of the participants. In addition to familiarity judgments, we collected concreteness judgments for individual senses, as well as word familiarity, and word concreteness judgments. The procedure we applied in collecting senses enabled us to estimate not only the number of senses, but their frequencies, that is, proportions of individual senses, as well. Based on these proportions we derived Information Theory measures – entropy and redundancy of the sense probability distribution. The collected measures will be the baseline for further research on the processing of polysemous words.

The results of the norming study revealed that the number of senses listed by the participants was much larger than the number of senses listed in the dictionary. However, after categorizing the senses listed by the participants according to the dictionary, the direction of this difference changed. The number of dictionary senses was larger than the number of dictionary senses appearing in the participants' descriptions. This finding was in accordance with the results of the studies conducted in English and Chinese (Gernsbacher, 1984; Lin & Ahrens, 2005). However, the number of dictionary senses that were rated as familiar by the participants was larger than the number of dictionary senses listed by the participants. A similar tendency was observed with the raw senses listed by the participants – although they listed two or three senses on average, they were familiar with most of

the senses collected in the study. This finding was in accordance with the assumption that participants are not able to list all of the familiar senses in a short period of time (Azuma, 1996). In spite of the differences in the number of senses that was estimated with various techniques, a positive correlation was observed among all of the collected measures. As expected, we observed a high correlation between sense frequencies and sense familiarity ratings. However, there were senses listed by a small number of participants only but at the same time rated as highly familiar.

The words selected for this study were generally rated by the participants as very familiar and mostly concrete. However, we noticed that word familiarity and word concreteness were related to the familiarity/concreteness of the dominant sense. This finding pointed to the fact that during the process of making a judgment on certain aspects of the whole word, participants were mostly relying on the dominant sense. In addition to this, we observed that concrete senses were more frequently listed,and were rated as more familiar.

Future research will be aiming at examining the relation among the collected measures in more detail. We believe that an understanding of the nature of these relations would contribute to understanding the way word senses are represented and processed. On the other hand, in further research we will explore the way entropy affects the processing of polysemous words. We will be particularly interested in the effects of the balance of sense probabilities.

References

Armstrong, B.C., Tokowicz, N., & Plaut, D.C. (2012). eDom: Norming software and relative meaning frequency norms for 544 homonyms. *Behavioral Research Methods, 44(4)*, 1015–1027.

Azuma, T. (1996). Familiarity and relatedness of word meanings: Ratings for 110 homographs. *Behavior Research Methods, Instruments, and Computers, 28(1)*, 109–124.

Azuma, T. and Van Orden, G. C. (1997). Why SAFE is better than FAST: The relatedness of a word's meanings affects lexical decision times. *Journal of Memory and Language, 36*, 484–504.

Baayen, R. H., Feldman, L. F. and Schreuder, R. (2006). Morphological influences on the recognition of monosyllabic monomorphemic words. *Journal of Memory and Language, 53*, 496–512.

Baayen, R. H., Milin, P., Filipovic Đurđevic, D., Hendrix, P. and Marelli, M. (2011), An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological Review, 118*, 438–482.

Balling, L. and Baayen, R.H. (2012) Probability and surprisal in auditory comprehension of morphologically complex words. *Cognition, 125*, 80–106.

Beretta, A., Fiorentino, R., Poeppel, D. (2005).The effects of homonymy and polysemy on lexical access: an MEG study. *Cognitive Brain Research, 24(1)*, 57–65.

Borowsky, R. and Masson, M. E. J. (1996). Semantic ambiguity effects in word identification. *Journal of Experimental Psychology: Learning,Memory, and Cognition, 22*, 63–85.

Connell, L., & Lynott, D. (2012). Strength of perceptual experience predicts word processing performance better than concreteness or imageability. *Cognition, 125*, 452–465.

Duffy, S.A., Morris, R.K., Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language, 27*, 429–446.

Ferraro, F. R. and Kellas, G. (1990). Normative data for number of word meanings. *Behavior Research Methods, Instruments, and Computers, 22(6)*, 491–498.

Filipović Đurđević, D., Đurđević, Đ., & Kostić, A. (2009). Vector based semantic analysis reveals absence of competition among related senses. *Psihologija, 42*(1), 95–106.

Forster, K. I., and Bednall, E. S. (1976). Terminating and exhaustive search in lexical access. *Memory and Cognition, 4*, 53–61.

Gawlick-Grendell, L. A., and Woltz, D. J. (1994). Meaning dominance norms for 120 homographs. *Behavior Research Methods, Instruments, and Computers, 26(1)*, 5–25.

Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General, 113*, 256–281.

Gilhooly, K.J. and Logie, R.H. (1980a). Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. *Behavior Research Methods and Instrumentation, 12(4)*, 395–427.

Gilhooly, K.J. and Logie, R.H. (1980b). Meaning-dependent ratings of imagery, age of acquisition, familiarity, and concreteness for 387 ambiguous words. *Behavior Research Methods and Instrumentation, 12(4)*, 428–450.

Hino, Y., and Lupker, S. J. (1996). The effects of polysemy in lexical decision and naming: An alternative to lexical access accounts. *Journal of Experimental Psychology: Human Perception and Performance, 22*, 1331–1356.

Jastrzembski, J. E. (1981). Multiple meanings, number of related mean-ings, frequency of occurrence, and the lexicon. *Cognitive Psychology, 13*, 278–305.

Jastrzembski, J. E., & Stanners, R. F. (1975). Multiple word meanings and lexical search speed. *Journal of Verbal Learning and Verbal Behavior, 14*, 534–537

Kellas, G., Ferraro, F. R., and Simpson, G. B. (1988). Lexical ambiguity and the timecourse of attentional allocation in word recognition. *Journal ofExperimental Psychology: Human Perception and Performance, 14*, 601–609.

Klepousniotou, E., & Baum, S.R. (2007). Disambiguating the ambiguity advantage in word recognition: an advantage for polysemous but not homonymous words. *Journal of Neurolinguistics, 20*, 1–24.

Klepousniotou, E., Pike, G.B., Steinhauer, K., Gracco, V. (2012). Not all ambiguous words are created equal: an EEG investigation

of homonymy and polysemy. *Brain and Language, 123(1)*, 11–21.

Klepousniotou, E., Titone, D., Romero, C. (2008). Making sense of word senses: the comprehension of polysemy depends on sense overlap. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34(6)*, 1534–1543.

Kostić, Đ. (1999). *Frekvencijski rečnik savremenog srpskog jezika. Tom I – VII*. Institut za eksperimentalnu fonetiku i patologiju govora, Beograd i Laboratorija za eksperimentalnu psihologiju Filozofskog fakulteta u Beogradu.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The Latent Semantic Analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review, 104*, 211–240.

Lund, K., & Burgess, C. (1997). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments & Computers, 28*, 203–208.

McDonald, S. (2000). *Environmental determinants of lexical processing effort*. PhD dissertation, University of Edinburgh.

Lynott, D., & Connell, L. (2009). Modality exclusivity norms for 423 object properties. *Behavior Research Methods, 41*, 558–564.

Lynott, D., & Connell, L. (2010). Embodied conceptual combination. *Frontiers in Psychology, 1*, 212.

Lynott, D., & Connell, L. (2013). Modality exclusivity norms for 400 nouns: The relationship between perceptual experience and surface word form. *Behavior Research Methods, 45*, 516–526.

Lin, C. C. and Ahrens, K. (2005). How Many Meanings Does A Word Have? Meaning Estimation in Chinese and English. In J. W. Minett and W. S-Y. Wang, editors, *Language Acquisition, Change and Emergence—Essays in Evolutionary Linguistics*, City University of Hong Kong Press, Hong Kong.

MacKay, D. J. C. (2003). *Information Theory, Inference, and Learning Algorithms*. Cambridge, UK: Cambridge University Press.

Milin, P., Filipović Đurđević, D., Kostić, A. & Moscoso del Prado Martìn, F. (2009). The simultaneous effects of inflectional

paradigms and classes on lexical recognition: Evidence from Serbian. *Journal of Memory and Language, 60*(1), 50–64.

Millis, M. L., and Button, S. B. (1989). The effect of polysemy on lexical decision time: Now you see it, now you don't. *Memory and Cognition, 17*,141–147.

Moscoso del Prado Martin, F., Kostić, A. and Baayen, H. (2004). Putting the bits together: An information-theoretical perspective on morphological processing. *Cognition, 94*, 1–18.

Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology, 76*(1p2), 1–25.

Paivio, A. (1986). *Mental representations: A dual coding approach*. New York: Oxford University Press.

Rečnik srpskohrvatskoga književnog jezika, 6 vols. Novi Sad: Matica Srpska; Zagreb: Matica Hrvatska, 1967-76.

Rodd, J. M., Gaskell, M. G., and Marslen-Wilson, W. D. (2002). Making sense of semantic ambiguity: Semantic competition in lexical access. *Journal of Memory and Language, 46*, 245–266.

Rubenstein, H., Garfield, L., and Millikan, J. A. (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior, 9*, 487–494.

Schütze, H. (1998). Automatic Word Sense Discrimination. *Computational Linguistics, 24* (1), 97–123

Schwanenflugel, P. (1991). Why are abstract concepts hard to understand.In P. J. Schwanenflugel (Ed.), *The Psychology of word meanings* (pp. 223–250). Hillsdale, NJ: Erlbaum.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal, XXVII*: 379–423.

Simpson, G.B. (1994). Context and the processing of ambiguous words. In: Gernsbacher, M. A. (editor), *Handbook of psycholinguistics* (pp. 359-374). San Diego (CA): Academic Press.

Swinney, D. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior, 18*, 645–659.

Tabak, W., Schreuder, R. and Baayen, R. H. (2005). Lexical statistics and lexical processing: semantic density, information complexity, sex, and irregularity in Dutch. In M. Reis and S. Kepser (eds), *Linguistic Evidence* (pp. 529–555). Mouton.

Twilley, L. C., Dixon, P., Taylor, D., and Clark, K. (1994). University of Alberta norms of relative meaning frequency for 566 homographs. *Memory and Cognition, 22(1)*, 111–126.

Whitney, C., Jefferies, E., Kircher, T. (2011). Heterogeneity of the left temporal lobe in semantic representation and control: Priming multiple vs. single meanings of ambiguous words. *Cerebral Cortex, 21*, 831–844.

Wurm, L.H., Ernestus, M., Schreuder, R., and Baayen, R. H. (2006) Dynamics of the Auditory Comprehension of Prefixed Words: Cohort Entropies and Conditional Root Uniqueness Points, *The Mental Lexicon*, 1, 125–146.

Appendix 1

Table of familiarity (left) and concreteness (right) ratings of a word, ratings averaged across individual senses listed in the dictionary *Rečnik Matice srpske* and ratings averaged across senses listed by the participants.

	Familiarity			Concreteness		
	word	mean senses, dictionary	mean senses, participants	word	mean senses, dictionary	mean senses, participants
album	6.56 (0.62)	5.98 (1.07)	5.41 (1.66)	6.52 (0.9)	6.2 (0.73)	6.35 (0.94)
balkon	6.19 (1.17)	4.77 (0.39)	6.48 (0.59)	6.95 (0.22)	6.48 (0.07)	5.79 (1.04)
balon	6.84 (0.37)	5.09 (0.49)	5.1 (0.84)	6.65 (0.57)	6.03 (1.1)	5.71 (1.44)
berba	5.67 (1.28)	5.76 (1.17)	4.55 (1.39)	5.29 (1.68)	3.71 (1.5)	5.02 (1.35)
blok	6.56 (0.7)	5.54 (0.66)	5.1 (1.81)	5.52 (1.47)	5.19 (1.35)	5.28 (1.54)
boks	5.82 (1.29)	4.85 (2.13)	4.67 (0.21)	5.7 (1.06)	5.37 (0.74)	5.25 (1.47)
brada	6.56 (0.86)	4.78 (0.58)	4.9 (0.29)	6.7 (0.56)	5.83 (1)	4.68 (1.64)
ćelija	6 (1.25)	5.47 (0.21)	4.2 (2.02)	4.24 (1.92)	5.31 (1.26)	4.23 (1.74)
centar	6.82 (0.53)	5.46 (1.26)	5.17 (1.48)	3.7 (1.33)	3.97 (1.37)	4.05 (1.89)
ciklus	5.59 (0.94)	5.24 (1.21)	6.74 (0.88)	1.96 (1.19)	4 (0.07)	3.8 (1.79)
članak	6.11 (1.05)	5.32 (1.09)	4.47 (0.67)	5.87 (1.25)	5.52 (0.87)	5.91 (1.13)
crevo	6.63 (0.68)	5.27 (0.29)	5.37 (0.2)	6.39 (0.89)	5.1 (1.69)	5.35 (1.53)
čvor	6.5 (0.71)	4.93 (1.76)	4.07 (0.22)	5.78 (1.04)	4.41 (2.2)	4.55 (1.58)
đavo	6.42 (0.84)	5.4 (1.26)	5.34 (1.37)	2.43 (1.66)	2.47 (0.64)	3.42 (2.19)
dinar	6.89 (0.32)	5.28 (0.24)	6.93 (0.21)	6.38 (0.92)	5.25 (0.07)	4.49 (1.92)

	Familiarity			Concreteness		
	word	mean senses, dictionary	mean senses, participants	word	mean senses, dictionary	mean senses, participants
disk	6.42 (0.84)	5.17 (0.61)	4.09 (1.15)	6.43 (0.84)	5.76 (0.8)	5.75 (1.53)
doktor	6.72 (0.67)	5.41 (0)	6.67 (1.07)	5.96 (1.15)	5.15 (1.63)	4.03 (1.51)
dugme	6.53 (0.84)	5.15 (1.45)	4.94 (0)	6.96 (0.21)	5.84 (1.24)	5.08 (1.51)
dvojka	6.71 (0.85)	4.62 (0.5)	3.17 (1.28)	4.04 (2.06)	5.07 (0.54)	4.48 (1.71)
figura	6.33 (0.84)	5.89 (1.26)	5.15 (1.35)	5.57 (1.63)	4.85 (1.45)	4.99 (1.53)
forma	6 (1.05)	5.36 (1.33)	5.1 (1)	2.65 (1.53)	4.04 (1.73)	4.33 (1.56)
gluma	6.47 (0.8)	6.26 (1.63)	5.86 (0.56)	2.57 (1.2)	4.46 (0.64)	4.14 (1.86)
govor	6.74 (0.56)	5.6 (1.27)	5.47 (0.93)	4.33 (1.83)	4.87 (0.78)	4.32 (1.63)
građa	6.06 (1.25)	5.05 (1.54)	5.46 (0.93)	4.13 (1.6)	5.17 (1.02)	4.01 (1.57)
grudi	6.42 (0.77)	4.93 (1.69)	5.5 (0.78)	6.48 (0.9)	5.15 (1.83)	4.35 (1.16)
grupa	6.47 (0.62)	5.37 (1.07)	5.61 (0.43)	4.96 (1.89)	4.2 (1.31)	4.52 (1.84)
guma	6.58 (0.9)	5.15 (1.52)	6.17 (1.81)	6.76 (0.54)	6.29 (0.82)	5.8 (1.59)
igrač	6.41 (1.18)	5.23 (0.96)	4.84 (0.49)	6.3 (1.15)	6 (0.8)	5.14 (1.32)
izbor	6.56 (0.63)	5.93 (0.81)	6.68 (0.72)	2.29 (1.49)	3.38 (0.32)	3.76 (1.74)
izlet	6.06 (1.09)	4.98 (0.33)	4.92 (0)	4.13 (1.52)	4.24 (0.54)	4.69 (1.61)
izraz	6.25 (1)	5.91 (0.48)	6.3 (0.78)	2.3 (1.18)	4.81 (0.91)	4.65 (1.93)
izvor	6.31 (1.3)	5.55 (0.79)	5.57 (0.64)	5.04 (1.3)	3.97 (1.48)	4.35 (1.9)
jezik	6.88 (0.34)	5.67 (0)	4.22 (0.64)	5.24 (1.79)	4.01 (1.4)	4.56 (1.86)
kanal	6.06 (1.25)	5.53 (1.69)	6.32 (0.69)	5.22 (1.57)	5.54 (2.07)	4.82 (1.62)

	Familiarity			Concreteness		
	word	mean senses, dictionary	mean senses, participants	word	mean senses, dictionary	mean senses, participants
kapak	6.24 (1.03)	3.78 (0)	3.37 (1.56)	6.3 (1.06)	5.62 (0.84)	5.33 (1.17)
ključ	6.83 (0.38)	5 (0)	3.75 (0)	6.57 (0.93)	4.32 (1.37)	4.1 (1.22)
kljun	5.88 (1.54)	3.98 (0.23)	3.74 (0)	6.74 (0.69)	5.19 (1.02)	5.54 (1.36)
klub	6.68 (0.58)	5.63 (1.11)	6.64 (0.87)	4.91 (1.7)	4.76 (1.82)	4.9 (1.59)
klupa	6.59 (0.71)	5.01 (1.43)	3.73 (1.29)	6.81 (0.51)	5.35 (1.2)	5.19 (1.31)
kolač	6.89 (0.32)	5.48 (0)	3.56 (1.24)	6.91 (0.29)	5.49 (1.21)	5.44 (1.25)
komad	6.47 (0.62)	5.75 (1.17)	6.1 (1.12)	5.26 (1.63)	5.35 (0.99)	5.75 (1.59)
komora	5.29 (1.76)	4.46 (1.12)	3.42 (1.4)	4.9 (1.64)	4.52 (1.23)	5.03 (1.56)
korak	6.65 (0.86)	5.6 (1.38)	5.76 (0)	5.48 (1.03)	4.12 (1.77)	3.62 (1.25)
koren	6.32 (0.89)	5.39 (0.53)	5.54 (0)	5.57 (1.47)	4.44 (1.59)	3.77 (1.68)
korica	6.28 (1.13)	4.51 (0.7)	5.06 (1.25)	6.57 (0.51)	5.44 (1.21)	5.83 (1.48)
koža	6.84 (0.5)	5.44 (0.46)	5.7 (0)	6.81 (0.51)	5.28 (1.97)	5.3 (1.23)
krug	6.82 (0.53)	5.14 (0.87)	5.34 (0.59)	4.65 (1.77)	4.38 (1.28)	4.49 (1.85)
kruna	6.31 (0.95)	5.16 (1.59)	4.58 (0.43)	6.43 (0.79)	4.85 (1.3)	4.91 (1.71)
krzno	6.32 (1.11)	5.36 (1.97)	6.38 (0.29)	6.81 (0.4)	6.48 (0.53)	5.34 (1.24)
kurs	6 (1.19)	5.74 (0.9)	5.69 (1.89)	2.57 (1.57)	3.41 (0.79)	3.44 (1.83)
lanac	6.63 (0.68)	5.6 (0.71)	4.44 (0.34)	6.22 (1.17)	4.23 (1.28)	4.49 (1.52)
linija	6.74 (0.65)	5.1 (1.91)	5.28 (0.63)	4.81 (1.47)	4 (1.53)	4.15 (1.38)
list	6.65 (0.79)	5.39 (0.42)	4.76 (0)	6.65 (0.65)	5.64 (0.74)	5.43 (1.14)

	Familiarity			Concreteness		
	word	mean senses, dictionary	mean senses, participants	word	mean senses, dictionary	mean senses, participants
lopta	6.79 (0.54)	5.04 (1.6)	6.05 (0.85)	6.57 (1.08)	5.82 (0.79)	4.94 (1.51)
loza	6 (1.29)	4.79 (1.12)	4.72 (0.34)	4.57 (1.75)	5.03 (1.61)	4.58 (1.29)
marka	6.63 (0.6)	5.3 (0.84)	4.75 (0.69)	5.26 (1.66)	4.49 (1.73)	4.62 (1.45)
masa	6.12 (1.17)	5.45 (1.02)	5.39 (0.42)	3.9 (1.79)	4.74 (1.27)	4.74 (1.5)
matica	5 (1.6)	4.51 (1.79)	4.62 (0.78)	4.43 (1.63)	4.77 (0.99)	4.98 (1.91)
mehur	5.67 (1.33)	4.44 (1.83)	4.71 (1.72)	6.05 (1.36)	5.63 (0.36)	5.38 (1.88)
mera	5.72 (1.32)	5.6 (1.95)	5.03 (1.14)	2.83 (1.7)	3.12 (0.98)	3.72 (1.69)
metar	6.53 (0.74)	5.22 (0)	5.38 (0.64)	4.57 (2.2)	3.95 (1.57)	4.5 (1.55)
minut	6.79 (0.54)	6.14 (0.39)	6.5 (0.29)	2.57 (1.5)	4.02 (1.09)	4.02 (2.11)
miris	6.44 (1.29)	5.85 (0.24)	6.46 (0.34)	4.76 (1.67)	5.28 (2.17)	4.61 (1.43)
model	6.44 (0.78)	5.97 (0.96)	5.63 (0.72)	4.71 (1.31)	5.43 (1.25)	5.43 (1.33)
momak	6.82 (0.73)	5.44 (0.24)	5.24 (0.29)	6.43 (0.98)	5.45 (0.7)	5.73 (1.36)
most	6.71 (0.77)	5.06 (0.24)	5.24 (1.04)	6.22 (1.28)	4.87 (1.78)	4.58 (1.78)
motiv	6.26 (0.99)	6.2 (0.33)	6.41 (0.29)	1.57 (1.21)	4.43 (1.83)	4.03 (2.05)
mreža	6.47 (1.07)	5.44 (1.15)	5.83 (0.66)	5.81 (1.54)	4.91 (1.67)	4.76 (1.65)
obim	6 (1.12)	5.56 (1.04)	6.24 (0.42)	3.6 (2.09)	3.83 (1.52)	4.65 (1.55)
oblast	5.82 (1.47)	5.45 (0.56)	5.2 (0.53)	3.7 (1.38)	4 (1.48)	4.7 (1.81)
oblik	6.18 (1.29)	5.71 (0.58)	5.26 (1.11)	4.05 (1.83)	4.45 (1.07)	4.48 (1.8)
obrada	6.05 (1.13)	5.61 (1.51)	5.98 (1.79)	2.67 (1.43)	4.13 (0.98)	4.65 (1.72)

NUMBER, RELATIVE FREQUENCY, ENTROPY, REDUNDANCY …

	Familiarity			Concreteness		
	word	mean senses, dictionary	mean senses, participants	word	mean senses, dictionary	mean senses, participants
obruč	5.24 (1.35)	4.86 (1.49)	4.54 (1.59)	6.13 (1.01)	3.96 (1.7)	4.89 (1.53)
ocena	6.78 (0.43)	6.11 (1.15)	6.39 (0.85)	4 (1.83)	4.05 (1.36)	3.42 (1.99)
oklop	5.75 (1.77)	5.28 (0.97)	5.75 (0.29)	6.52 (0.98)	5.29 (1.61)	5.8 (1.31)
organ	6.53 (0.51)	6 (0.24)	4.58 (0.21)	5.38 (1.63)	4.6 (1.06)	4.05 (1.73)
osnova	6.22 (0.94)	5.23 (1.21)	5.28 (1.4)	2.81 (1.25)	4 (1.17)	3.94 (1.61)
patent	4.94 (1.69)	5.12 (1.21)	4.91 (1.35)	4.29 (1.74)	4.48 (0.95)	5.35 (1.35)
pesak	6.29 (0.92)	5.49 (0.64)	6.15 (1.85)	6.65 (0.57)	6.43 (0.53)	6.13 (1.03)
pešak	6.65 (0.7)	4.84 (1.59)	5.57 (0)	6.61 (0.78)	6.38 (0.19)	5.69 (1.2)
pisak	3.71 (1.69)	4.71 (1.7)	3.43 (2.1)	5.19 (1.63)	5.8 (0.49)	6.1 (0.99)
pismo	6.71 (0.77)	5.85 (1.33)	4.84 (0.7)	6.48 (1.24)	4.96 (0.97)	5.51 (1.67)
platno	6.39 (0.85)	5.27 (0.29)	4.17 (1.49)	6.33 (0.8)	5.48 (1.48)	5.57 (1.59)
ploča	6.53 (0.77)	5.49 (0.8)	4.18 (1.07)	6.48 (0.85)	5.21 (0.88)	5.87 (1.29)
plod	6.29 (1.05)	6.05 (1.01)	6.42 (0.22)	6.29 (1.06)	5.03 (2.36)	5.42 (1.18)
pojas	6.35 (1.17)	5.33 (1.37)	4.04 (1.58)	5.86 (1.35)	4.17 (1.52)	5.23 (1.54)
pojava	6.33 (0.91)	5.93 (1.09)	4.83 (1.98)	2.52 (1.44)	3.72 (1.44)	4.74 (1.75)
pokret	6.53 (0.87)	5.74 (1)	5.06 (1.05)	4.39 (1.78)	4.36 (1.87)	4.36 (1.62)
polet	5.28 (1.18)	4.05 (1.42)	5.68 (2.08)	1.83 (1.37)	3.44 (1.26)	3.52 (1.85)
posao	6.88 (0.33)	5.89 (1.18)	5.66 (0.53)	3.48 (1.75)	3.87 (0.62)	3.84 (1.94)
poskok	5.67 (1.37)	4.06 (0)	6.19 (1.92)	6.48 (0.98)	6.24 (1.01)	5.84 (0.82)

	Familiarity			Concreteness		
	word	mean senses, dictionary	mean senses, participants	word	mean senses, dictionary	mean senses, participants
pošta	6.81 (0.4)	5.64 (0.43)	4.42 (1.86)	5.86 (1.15)	4.73 (1.62)	5.03 (1.51)
potez	5.67 (1.46)	5.53 (1.28)	4.75 (2.03)	3.3 (1.58)	4.13 (0.94)	4.56 (1.82)
poziv	6.37 (0.9)	5.78 (0.31)	5.61 (0.71)	3.57 (1.95)	4.22 (1.67)	4.16 (1.79)
prašak	6.33 (0.84)	5.1 (0.64)	3.95 (1.14)	6.48 (0.73)	4.13 (1.55)	5.98 (1.06)
pravac	6.67 (0.59)	5.27 (0.31)	4.97 (1.43)	2.57 (1.31)	3.61 (1.34)	3.43 (1.83)
prenos	6.06 (1.2)	5.55 (1.02)	5.78 (0.8)	2.7 (1.33)	4.51 (1.16)	3.93 (2.01)
prilog	6.18 (1.13)	6.49 (0.73)	6.03 (1.13)	3.48 (1.75)	4.86 (0.73)	4.43 (1.54)
profil	6.21 (1.03)	4.7 (0)	5.11 (0.28)	4.91 (1.65)	4.73 (1.17)	3.88 (1.72)
pruga	6.22 (1.35)	5.09 (0.66)	3.74 (0.39)	6.67 (0.73)	4.72 (1.38)	5.64 (1.14)
račun	6.71 (0.47)	5.98 (1.09)	4.91 (1.77)	5.7 (1.26)	3.67 (1.2)	3.58 (1.69)
radnja	6.67 (0.49)	5.95 (0.75)	4.5 (0.95)	3.52 (1.9)	4.77 (1.03)	4.58 (1.5)
rebro	6.12 (1.36)	4.93 (0.49)	3.39 (0)	6.57 (1.16)	5.3 (0.53)	5.89 (0.9)
salon	6.32 (1.16)	5.53 (0.86)	5.53 (1.32)	6.1 (0.83)	5.49 (1.38)	6.42 (0.83)
samica	5.67 (1.33)	4.96 (0.75)	4.67 (1.92)	4.91 (1.68)	5.88 (0.89)	5.49 (1.38)
sastav	6.18 (1.13)	5.53 (0.5)	5.49 (0.9)	3.57 (1.85)	4.52 (1.04)	4.66 (1.87)
savet	6.63 (0.68)	5.85 (0.21)	5.43 (0)	2.48 (1.72)	4.08 (0.38)	4.02 (1.75)
scena	6.26 (0.93)	5.75 (1.34)	6.45 (0.58)	4.33 (1.68)	5.19 (0.84)	4.69 (1.62)
servis	5.59 (1.62)	6.04 (1.66)	5.64 (1.85)	4.95 (1.66)	5.31 (1.18)	5.06 (1.69)
sfera	4.89 (1.64)	4.92 (1.89)	5.66 (1.33)	3.29 (1.71)	4.03 (1.39)	3.71 (1.82)

	Familiarity			Concreteness		
	word	mean senses, dictionary	mean senses, participants	word	mean senses, dictionary	mean senses, participants
signal	5.59 (1.42)	5.85 (0.84)	5.65 (1)	4.14 (1.59)	4.35 (1.78)	5.34 (1.6)
sila	6.53 (0.72)	5.65 (1.53)	5.16 (1.02)	2.26 (1.39)	3.59 (0.74)	3.3 (1.84)
sistem	6.42 (0.84)	5.41 (1.5)	4.94 (1.93)	2.35 (1.23)	3.62 (1.25)	3.7 (1.97)
skakač	5.95 (1.08)	5.56 (0.58)	5 (1.95)	5.87 (0.97)	6.19 (0.38)	6.56 (0.83)
skok	6.5 (0.82)	5.25 (1.99)	5.08 (0.94)	4.67 (1.62)	3.99 (1.67)	4.65 (1.24)
škola	6.95 (0.23)	5.72 (1.19)	5.78 (1.33)	6.1 (0.94)	4.32 (1.12)	3.67 (1.84)
sloj	5.74 (1.33)	5.34 (1.61)	6.65 (0.73)	4.09 (1.7)	4.78 (2.37)	4.54 (1.35)
sluh	6.47 (1.07)	5.79 (0.21)	5.5 (0)	3.05 (1.77)	3.43 (1.2)	3.51 (2.03)
smer	5.83 (1.15)	6.05 (1.09)	4.97 (0.69)	3.43 (1.89)	3.02 (1.25)	4.06 (1.8)
snimak	6.12 (1.27)	5.33 (1.36)	5.62 (0.8)	5.52 (1.36)	5.53 (1.45)	5.84 (1.06)
stav	6.5 (0.62)	5.74 (1.33)	5.14 (0.77)	1.52 (0.85)	3.93 (1.12)	4.36 (1.57)
stena	6 (1.14)	4.79 (0.99)	4.14 (1.1)	6.81 (0.4)	4.92 (2.06)	3.72 (1.55)
stepen	5.94 (1.26)	5.92 (1.09)	5.21 (0.9)	2.35 (1.72)	3.64 (0.91)	3.3 (1.97)
stopa	5.94 (1.11)	4.85 (1.96)	4.21 (1.56)	4.52 (1.9)	4.41 (1.22)	4.56 (1.59)
struja	6.44 (1.15)	5.38 (0.22)	5.1 (1.35)	4.86 (1.65)	4.13 (1.1)	3.34 (1.86)
struk	6.22 (1)	5.03 (1.86)	4 (1.82)	6.04 (1.19)	5.34 (0.65)	5.8 (1.23)
tabla	6.82 (0.39)	5.22 (0.29)	4.3 (1.43)	6.78 (0.52)	5.2 (1.46)	5.56 (1.32)
tačka	6.82 (0.73)	5.64 (0.58)	5.39 (1.62)	4.57 (2.01)	4.2 (1.22)	4.14 (1.84)
talas	6.47 (0.7)	5.6 (0.29)	5.54 (0)	5.43 (1.63)	4.83 (1.41)	4.65 (1.7)

	Familiarity			Concreteness		
	word	mean senses, dictionary	mean senses, participants	word	mean senses, dictionary	mean senses, participants
tekst	6.65 (0.7)	6.32 (0.47)	5.79 (0.22)	6.1 (1.3)	5.86 (0.37)	4.92 (1.41)
termin	6 (1.46)	6.16 (0.96)	5.63 (0.37)	2.57 (1.75)	3.49 (0.73)	3.56 (1.91)
teza	4.71 (1.76)	5.69 (1.46)	4.59 (1.28)	1.96 (1.46)	3.71 (1.02)	3.66 (1.6)
tezga	6.16 (1.12)	6.03 (0)	6.5 (0.31)	6.38 (0.97)	4.88 (2.65)	4.6 (1.38)
traka	6.63 (0.68)	5.2 (1.07)	4.39 (0.47)	6.29 (0.72)	5.18 (1.71)	6.09 (1.06)
ugao	6.13 (1.26)	6.01 (0.74)	5.95 (0.61)	4.39 (1.41)	5.23 (1.43)	5.17 (1.58)
ukus	6.63 (0.83)	6.02 (1.05)	6.57 (1.05)	4.09 (1.78)	4.45 (1.95)	3.58 (1.75)
uslov	5.84 (1.21)	5.96 (0.35)	6.83 (0.2)	2.05 (1.43)	3.39 (0.06)	3.32 (1.93)
uspon	5.94 (1.39)	5.49 (0.53)	5.07 (0.93)	4.1 (1.92)	4.61 (1.12)	4.13 (1.69)
vatra	6.74 (0.56)	4.99 (0.82)	4.5 (0.49)	6.24 (0.94)	4.02 (1.38)	4.21 (1.39)
venac	5.94 (1.26)	5.15 (0.54)	4.15 (0.21)	6.39 (1.03)	4.87 (0.85)	4.87 (1.7)
veza	6.61 (0.5)	5.51 (1.12)	5.9 (1.97)	3.14 (1.31)	4.16 (1.07)	3.63 (1.43)
visina	6.71 (0.77)	5.56 (0)	5.69 (1.36)	4 (1.62)	4.11 (1)	4.47 (1.65)
vrat	6.44 (1.15)	4.74 (0)	6.11 (0)	7 (0)	6.43 (0.54)	5.13 (1.87)
vreća	6.39 (0.98)	4.54 (0.96)	4 (0.29)	6.52 (0.73)	4.11 (2.05)	4.9 (1.17)
zakon	6.32 (1.16)	5.65 (0.5)	5.01 (1)	3.05 (1.94)	2.81 (0.88)	3.01 (1.52)
zglob	6.35 (0.86)	4.33 (0.23)	3.16 (0)	6.52 (0.85)	4 (1.92)	5.66 (1.4)
žica	6.33 (0.97)	5.17 (1.54)	3.98 (1.78)	6.62 (0.67)	4.38 (1.75)	5.05 (1.17)
znak	6.28 (0.83)	5.61 (1.15)	5.67 (1.92)	4.76 (1.87)	5.01 (0.95)	4.42 (1.63)

NUMBER, RELATIVE FREQUENCY, ENTROPY, REDUNDANCY …

	Familiarity			Concreteness		
	word	mean senses, dictionary	mean senses, participants	word	mean senses, dictionary	mean senses, participants
zrno	6.56 (0.62)	5.51 (0.44)	5.47 (0.43)	6.3 (0.88)	6.1 (0.62)	5.39 (1.56)
zvono	6.61 (0.85)	4.78 (1.43)	5.25 (0.21)	6.7 (1.06)	4.75 (1.86)	5.2 (1.18)

Appendix 2

Table of number of senses (N), entropy (H) and redundancy (T) of sense probability distribution based on the senses a) listed by the participants, b) listed by more than 10% of the participants, c) with a mean familiarity rating of 2 or higher, d) with a mode familiarity rating above 1, and e) with a median familiarity rating above 1.

	a) Senses listed by participants			b) Senses listed by more than 10% of participants			c) Senses with mean familiarity rating>1.9			d) Senses with mode familiarity rating >1			e) Senses with median familiarity rating >1		
	N	H	T	N	H	T	N	H	T	N	H	T	N	H	T
album	5	2.12	0.09	5	2.12	0.09	5	2.12	0.09	5	2.12	0.09	5	2.12	0.09
balkon	6	1.81	0.30	4	1.5	0.25	5	1.57	0.33	5	1.57	0.33	5	1.57	0.33
balon	12	3.13	0.13	10	3	0.1	12	3.13	0.13	10	2.94	0.12	12	3.13	0.13
berba	8	2.76	0.08	6	2.46	0.05	8	2.76	0.08	7	2.61	0.07	8	2.76	0.08
blok	16	3.54	0.11	14	3.43	0.1	16	3.54	0.11	16	3.54	0.11	16	3.54	0.11
boks	12	2.97	0.17	7	2.51	0.1	12	2.97	0.17	9	2.71	0.15	12	2.97	0.17
brada	7	2.00	0.29	5	1.77	0.24	7	2.00	0.29	5	1.70	0.27	7	2.00	0.29
ćelija	13	2.91	0.21	8	2.51	0.16	12	2.83	0.21	12	2.83	0.21	12	2.83	0.21
centar	18	3.70	0.11	11	3.29	0.05	17	3.65	0.11	17	3.65	0.11	17	3.65	0.11
ciklus	21	3.80	0.13	9	2.93	0.08	21	3.80	0.13	19	3.68	0.13	21	3.80	0.13
članak	8	2.12	0.29	5	1.79	0.23	9	2.01	0.36	9	2.01	0.36	9	2.01	0.36
crevo	11	2.51	0.27	6	2.05	0.21	12	2.51	0.30	12	2.51	0.30	12	2.51	0.3
čvor	17	3.53	0.14	10	2.96	0.11	16	3.46	0.13	16	3.46	0.13	16	3.46	0.13
đavo	25	4.19	0.10	12	3.42	0.05	25	4.14	0.11	23	4.09	0.10	24	4.14	0.1

	a) Senses listed by participants			b) Senses listed by more than 10% of participants			c) Senses with mean familiarity rating>1.9			d) Senses with mode familiarity rating >1			e) Senses with median familiarity rating >1		
	N	H	T	N	H	T	N	H	T	N	H	T	N	H	T
dinar	6	1.96	0.24	4	1.73	0.13	6	1.96	0.24	5	1.85	0.20	6	1.96	0.24
disk	13	3.06	0.17	9	2.74	0.14	13	3.06	0.17	10	2.73	0.18	13	3.06	0.17
doktor	4	1.60	0.20	3	1.46	0.08	4	1.60	0.20	4	1.60	0.20	4	1.60	0.2
dugme	6	2.24	0.13	6	2.24	0.13	6	2.24	0.13	5	2.02	0.13	6	2.24	0.13
dvojka	16	3.30	0.17	8	2.73	0.09	15	3.24	0.17	11	2.80	0.19	14	3.11	0.18
figura	14	3.39	0.11	11	3.22	0.07	14	3.39	0.11	14	3.39	0.11	14	3.39	0.11
forma	15	3.22	0.18	10	2.85	0.14	13	3.08	0.17	13	3.08	0.17	13	3.08	0.17
gluma	7	2.44	0.13	5	2.2	0.05	7	2.44	0.13	7	2.44	0.13	7	2.44	0.13
govor	17	3.39	0.17	9	2.84	0.11	17	3.33	0.19	17	3.33	0.19	17	3.33	0.19
građa	8	2.23	0.26	4	1.78	0.11	8	2.23	0.26	7	2.12	0.24	8	2.23	0.26
grudi	12	2.67	0.26	7	2.17	0.23	12	2.67	0.26	10	2.48	0.25	12	2.67	0.26
grupa	25	4.11	0.12	15	3.54	0.1	25	4.11	0.12	25	4.11	0.12	25	4.11	0.12
guma	15	3.18	0.19	11	2.96	0.14	15	3.18	0.19	15	3.18	0.19	15	3.18	0.19
igrač	10	2.61	0.22	6	2.21	0.14	10	2.61	0.22	10	2.61	0.22	10	2.61	0.22
izbor	16	3.58	0.10	8	2.89	0.04	16	3.58	0.10	16	3.58	0.10	16	3.58	0.1
izlet	11	2.97	0.14	8	2.65	0.12	11	2.97	0.14	11	2.97	0.14	11	2.97	0.14
izraz	10	3.10	0.07	10	3.1	0.07	10	3.10	0.07	10	3.10	0.07	10	3.10	0.07
izvor	10	2.85	0.14	7	2.57	0.09	10	2.85	0.14	10	2.85	0.14	10	2.85	0.14
jezik	16	3.24	0.19	8	2.61	0.13	16	3.24	0.19	14	3.09	0.19	15	3.17	0.19

	a) Senses listed by participants			b) Senses listed by more than 10% of participants			c) Senses with mean familiarity rating>1.9			d) Senses with mode familiarity rating >1			e) Senses with median familiarity rating >1		
	N	H	T	N	H	T	N	H	T	N	H	T	N	H	T
kanal	20	3.89	0.10	13	3.46	0.06	20	3.89	0.10	18	3.78	0.09	20	3.89	0.1
kapak	9	2.29	0.28	5	1.8	0.22	8	2.18	0.27	5	1.72	0.26	7	1.99	0.29
ključ	13	3.12	0.16	8	2.79	0.07	13	3.12	0.16	11	2.99	0.13	13	3.12	0.16
kljun	10	2.62	0.21	6	2.2	0.15	10	2.62	0.21	6	1.81	0.30	9	2.52	0.2
klub	13	3.04	0.18	8	2.66	0.11	13	3.04	0.18	13	3.04	0.18	13	3.04	0.18
klupa	11	2.68	0.22	5	2.06	0.11	10	2.59	0.22	10	2.59	0.22	10	2.59	0.22
kolač	9	2.71	0.14	7	2.5	0.11	9	2.71	0.14	8	2.49	0.17	9	2.71	0.14
komad	13	3.11	0.16	8	2.69	0.1	13	3.11	0.16	13	3.11	0.16	13	3.11	0.16
komora	20	3.67	0.15	7	2.63	0.06	20	3.67	0.15	17	3.42	0.16	20	3.67	0.15
korak	13	3.34	0.10	10	3.11	0.06	13	3.34	0.10	13	3.34	0.10	13	3.34	0.1
koren	16	3.39	0.15	12	3.17	0.12	16	3.39	0.15	15	3.33	0.15	16	3.39	0.15
korica	15	3.32	0.15	8	2.69	0.1	14	3.24	0.15	14	3.24	0.15	15	3.32	0.15
koža	15	3.13	0.20	7	2.47	0.12	15	3.13	0.20	14	3.06	0.20	15	3.13	0.2
krug	22	3.82	0.14	11	3.01	0.13	22	3.82	0.14	21	3.76	0.14	22	3.82	0.14
kruna	13	2.98	0.19	9	2.66	0.16	13	2.98	0.19	11	2.78	0.20	12	2.86	0.2
krzno	6	1.98	0.24	4	1.76	0.12	6	1.98	0.24	5	1.87	0.19	6	1.98	0.24
kurs	6	2.27	0.12	5	2.16	0.07	6	2.27	0.12	6	2.27	0.12	6	2.27	0.12
lanac	20	3.59	0.17	10	3.01	0.09	19	3.53	0.17	19	3.53	0.17	19	3.53	0.17
linija	34	4.48	0.12	14	3.61	0.05	35	4.48	0.13	34	4.45	0.13	35	4.48	0.13

NUMBER, RELATIVE FREQUENCY, ENTROPY, REDUNDANCY …

	a) Senses listed by participants			b) Senses listed by more than 10% of participants			c) Senses with mean familiarity rating>1.9			d) Senses with mode familiarity rating >1			e) Senses with median familiarity rating >1		
	N	H	T	N	H	T	N	H	T	N	H	T	N	H	T
list	12	3.05	0.15	9	2.84	0.1	12	3.05	0.15	11	2.99	0.14	12	3.05	0.15
lopta	11	2.88	0.17	7	2.56	0.09	11	2.88	0.17	10	2.81	0.16	11	2.88	0.17
loza	10	2.61	0.21	7	2.36	0.16	8	2.40	0.20	8	2.40	0.20	8	2.40	0.2
marka	10	2.61	0.21	8	2.46	0.18	10	2.61	0.21	9	2.54	0.20	10	2.61	0.21
masa	11	3.01	0.13	8	2.79	0.07	11	3.01	0.13	11	3.01	0.13	11	3.01	0.13
matica	12	3.06	0.15	8	2.74	0.09	11	2.99	0.14	9	2.78	0.12	10	2.86	0.14
mehur	12	3.06	0.15	7	2.6	0.07	11	2.98	0.14	10	2.79	0.16	11	2.98	0.14
mera	14	3.29	0.14	10	2.96	0.11	14	3.29	0.14	14	3.29	0.14	14	3.29	0.14
metar	9	2.52	0.21	7	2.3	0.18	9	2.52	0.21	8	2.35	0.22	8	2.35	0.22
minut	8	2.60	0.13	6	2.4	0.07	8	2.60	0.13	8	2.60	0.13	8	2.60	0.13
miris	8	2.60	0.13	7	2.5	0.11	8	2.60	0.13	8	2.60	0.13	8	2.60	0.13
model	14	3.61	0.05	13	3.56	0.04	14	3.61	0.05	14	3.61	0.05	14	3.61	0.05
momak	6	1.89	0.27	3	1.48	0.06	6	1.89	0.27	6	1.89	0.27	6	1.89	0.27
most	15	3.17	0.19	7	2.5	0.11	15	3.17	0.19	14	3.09	0.19	15	3.17	0.19
motiv	10	2.98	0.10	9	2.9	0.09	10	2.98	0.10	10	2.98	0.10	10	2.98	0.1
mreža	25	4.10	0.12	16	3.64	0.09	25	4.10	0.12	25	4.10	0.12	25	4.10	0.12
obim	8	2.47	0.18	5	2.09	0.1	8	2.47	0.18	8	2.47	0.18	8	2.47	0.18
oblast	11	3.11	0.10	9	2.96	0.07	11	3.11	0.10	11	3.11	0.10	11	3.11	0.1
oblik	13	2.97	0.20	6	2.09	0.19	13	2.97	0.20	13	2.97	0.20	13	2.97	0.2

	a) Senses listed by participants			b) Senses listed by more than 10% of participants			c) Senses with mean familiarity rating>1.9			d) Senses with mode familiarity rating >1			e) Senses with median familiarity rating >1		
	N	H	T	N	H	T	N	H	T	N	H	T	N	H	T
obrada	14	3.29	0.14	10	3.01	0.09	14	3.29	0.14	14	3.29	0.14	14	3.29	0.14
obruč	14	3.19	0.16	8	2.7	0.1	14	3.19	0.16	13	3.11	0.16	14	3.19	0.16
ocena	6	2.31	0.11	6	2.31	0.11	6	2.31	0.11	6	2.31	0.11	6	2.31	0.11
oklop	9	2.74	0.14	7	2.53	0.1	9	2.74	0.14	9	2.74	0.14	9	2.74	0.14
organ	5	1.84	0.21	4	1.71	0.15	5	1.84	0.21	5	1.84	0.21	5	1.84	0.21
osnova	11	3.04	0.12	8	2.77	0.08	12	2.95	0.18	11	2.95	0.15	11	2.95	0.15
patent	7	2.35	0.16	5	2.06	0.11	6	2.21	0.15	6	2.21	0.15	6	2.21	0.15
pesak	11	3.13	0.10	7	2.69	0.04	11	3.13	0.10	9	2.85	0.10	11	3.13	0.1
pešak	7	2.18	0.22	4	1.84	0.08	7	2.18	0.22	5	1.96	0.16	6	2.07	0.2
pisak	12	3.26	0.09	9	2.97	0.06	11	3.17	0.08	9	2.89	0.09	11	3.17	0.08
pismo	8	2.37	0.21	4	1.94	0.03	8	2.37	0.21	8	2.37	0.21	8	2.37	0.21
platno	8	2.11	0.30	3	1.52	0.04	8	2.11	0.30	7	2.00	0.29	7	2.00	0.29
ploča	18	3.45	0.17	12	3.07	0.14	18	3.45	0.17	17	3.36	0.18	18	3.45	0.17
plod	10	2.87	0.14	7	2.59	0.08	10	2.87	0.14	10	2.87	0.14	10	2.87	0.14
pojas	19	3.70	0.13	11	3.19	0.08	17	3.58	0.12	16	3.52	0.12	17	3.58	0.12
pojava	16	3.63	0.09	13	3.43	0.07	16	3.63	0.09	16	3.63	0.09	16	3.63	0.09
pokret	20	3.81	0.12	9	2.99	0.06	20	3.81	0.12	20	3.81	0.12	20	3.81	0.12
polet	12	2.94	0.18	6	2.28	0.12	9	2.63	0.17	7	2.40	0.14	8	2.52	0.16
posao	9	2.84	0.11	7	2.66	0.05	9	2.84	0.11	8	2.75	0.08	9	2.84	0.11

	a) Senses listed by participants			b) Senses listed by more than 10% of participants			c) Senses with mean familiarity rating>1.9			d) Senses with mode familiarity rating >1			e) Senses with median familiarity rating >1		
	N	H	T	N	H	T	N	H	T	N	H	T	N	H	T
poskok	7	1.96	0.30	3	1.42	0.1	5	1.70	0.27	5	1.70	0.27	5	1.70	0.27
pošta	9	2.71	0.15	6	2.41	0.07	9	2.71	0.15	9	2.71	0.15	9	2.71	0.15
potez	11	2.88	0.17	7	2.53	0.1	10	2.80	0.16	9	2.57	0.19	10	2.80	0.16
poziv	13	3.18	0.14	9	2.89	0.09	12	3.11	0.13	12	3.11	0.13	12	3.11	0.13
prašak	11	3.08	0.11	7	2.75	0.02	12	3.08	0.14	9	2.83	0.11	12	3.08	0.14
pravac	18	3.57	0.14	12	3.19	0.11	18	3.57	0.14	16	3.45	0.14	18	3.57	0.14
prenos	14	3.04	0.20	7	2.46	0.12	14	3.04	0.20	14	3.04	0.20	14	3.04	0.2
prilog	8	2.54	0.15	7	2.44	0.13	8	2.54	0.15	8	2.54	0.15	8	2.54	0.15
profil	15	3.21	0.18	10	2.82	0.15	15	3.21	0.18	13	3.02	0.18	14	3.10	0.19
pruga	9	2.53	0.20	6	2.21	0.14	9	2.53	0.20	9	2.53	0.20	9	2.53	0.2
račun	14	3.30	0.13	11	3.11	0.1	14	3.30	0.13	14	3.30	0.13	14	3.30	0.13
radnja	8	2.28	0.24	7	2.19	0.22	8	2.28	0.24	8	2.28	0.24	8	2.28	0.24
rebro	8	2.34	0.22	7	2.23	0.2	8	2.34	0.22	6	1.89	0.27	8	2.34	0.22
salon	11	2.92	0.16	8	2.64	0.12	11	2.83	0.18	11	2.83	0.18	11	2.83	0.18
samica	9	2.37	0.25	4	1.7	0.15	9	2.37	0.25	8	2.14	0.29	9	2.37	0.25
sastav	12	3.08	0.14	8	2.78	0.07	12	3.08	0.14	10	2.93	0.12	11	3.01	0.13
savet	11	2.91	0.16	9	2.73	0.14	10	2.74	0.17	10	2.74	0.17	10	2.74	0.17
scena	19	3.33	0.22	7	2.48	0.12	16	3.14	0.21	16	3.14	0.21	16	3.14	0.21
servis	10	2.70	0.19	6	2.35	0.09	10	2.70	0.19	10	2.70	0.19	10	2.70	0.19

	a) Senses listed by participants			b) Senses listed by more than 10% of participants			c) Senses with mean familiarity rating>1.9			d) Senses with mode familiarity rating >1			e) Senses with median familiarity rating >1		
	N	H	T	N	H	T	N	H	T	N	H	T	N	H	T
sfera	11	2.78	0.20	7	2.41	0.14	11	2.78	0.20	8	2.46	0.18	11	2.78	0.2
signal	16	3.83	0.04	14	3.7	0.03	15	3.74	0.04	15	3.74	0.04	15	3.74	0.04
sila	24	4.04	0.12	11	3.27	0.05	24	4.04	0.12	23	3.99	0.12	23	3.99	0.12
sistem	25	4.10	0.12	12	3.37	0.06	24	4.05	0.12	24	4.05	0.12	24	4.05	0.12
skakač	9	2.51	0.21	5	2.04	0.12	9	2.51	0.21	9	2.51	0.21	9	2.51	0.21
skok	14	3.31	0.13	8	2.73	0.09	13	3.22	0.13	13	3.22	0.13	13	3.22	0.13
škola	14	3.11	0.18	7	2.51	0.11	14	3.11	0.18	13	3.03	0.18	14	3.11	0.18
sloj	12	3.08	0.14	9	2.83	0.11	11	2.96	0.15	11	2.96	0.15	11	2.96	0.15
sluh	6	1.95	0.24	5	1.84	0.21	6	1.95	0.24	6	1.95	0.24	6	1.95	0.24
smer	13	3.30	0.11	10	3.04	0.08	13	3.30	0.11	13	3.30	0.11	13	3.30	0.11
snimak	12	3.02	0.16	6	2.38	0.08	12	3.02	0.16	12	3.02	0.16	12	3.02	0.16
stav	12	2.64	0.26	6	2.04	0.21	12	2.64	0.26	12	2.64	0.26	12	2.64	0.26
stena	17	3.43	0.16	6	2.38	0.08	17	3.43	0.16	14	3.18	0.16	17	3.43	0.16
stepen	13	3.06	0.17	7	2.57	0.08	11	2.86	0.17	11	2.86	0.17	11	2.86	0.17
stopa	10	2.78	0.16	7	2.53	0.1	9	2.70	0.15	8	2.62	0.13	9	2.70	0.15
struja	10	2.75	0.17	8	2.56	0.15	10	2.75	0.17	9	2.66	0.16	10	2.75	0.17
struk	9	2.48	0.22	7	2.25	0.2	8	2.37	0.21	8	2.37	0.21	8	2.37	0.21
tabla	16	3.28	0.18	8	2.64	0.12	15	3.21	0.18	14	3.13	0.18	15	3.21	0.18
tačka	15	3.34	0.14	11	3.08	0.11	15	3.34	0.14	15	3.34	0.14	15	3.34	0.14

	a) Senses listed by participants			b) Senses listed by more than 10% of participants			c) Senses with mean familiarity rating>1.9			d) Senses with mode familiarity rating >1			e) Senses with median familiarity rating >1		
	N	H	T	N	H	T	N	H	T	N	H	T	N	H	T
talas	19	3.75	0.12	14	3.47	0.09	19	3.75	0.12	18	3.70	0.11	19	3.75	0.12
tekst	6	2.26	0.12	5	2.12	0.09	6	2.26	0.12	6	2.26	0.12	6	2.26	0.12
termin	6	2.21	0.14	6	2.21	0.14	6	2.21	0.14	6	2.21	0.14	6	2.21	0.14
teza	7	2.46	0.12	6	2.33	0.1	7	2.46	0.12	7	2.46	0.12	7	2.46	0.12
tezga	6	2.09	0.19	5	1.97	0.15	6	2.09	0.19	6	2.09	0.19	6	2.09	0.19
traka	25	4.21	0.09	17	3.85	0.06	25	4.21	0.09	23	4.12	0.09	25	4.21	0.09
ugao	8	2.69	0.10	7	2.6	0.07	8	2.69	0.10	7	2.60	0.07	8	2.69	0.1
ukus	11	2.93	0.15	7	2.57	0.08	10	2.84	0.14	10	2.84	0.14	10	2.84	0.14
uslov	11	2.97	0.14	8	2.66	0.11	11	2.97	0.14	11	2.97	0.14	11	2.97	0.14
uspon	8	2.35	0.22	5	2.01	0.13	8	2.35	0.22	6	2.13	0.18	8	2.35	0.22
vatra	19	3.84	0.10	11	3.36	0.03	18	3.78	0.09	18	3.78	0.09	18	3.78	0.09
venac	15	3.37	0.14	9	2.94	0.07	15	3.37	0.14	14	3.30	0.13	15	3.37	0.14
veza	23	3.86	0.15	14	3.37	0.11	23	3.86	0.15	22	3.81	0.15	23	3.86	0.15
visina	17	3.52	0.14	8	2.81	0.06	17	3.52	0.14	17	3.52	0.14	17	3.52	0.14
vrat	13	2.70	0.27	5	1.63	0.3	13	2.70	0.27	13	2.70	0.27	13	2.70	0.27
vreća	12	2.95	0.18	6	2.27	0.12	10	2.69	0.19	9	2.58	0.19	11	2.80	0.19
zakon	8	2.54	0.15	7	2.45	0.13	8	2.54	0.15	8	2.54	0.15	8	2.54	0.15
zglob	7	2.14	0.24	5	1.88	0.19	7	2.14	0.24	5	1.63	0.30	7	2.14	0.24
žica	15	3.56	0.09	12	3.37	0.06	15	3.56	0.09	14	3.43	0.10	15	3.56	0.09

	a) Senses listed by participants			b) Senses listed by more than 10% of participants			c) Senses with mean familiarity rating>1.9			d) Senses with mode familiarity rating >1			e) Senses with median familiarity rating >1		
	N	H	T	N	H	T	N	H	T	N	H	T	N	H	T
znak	13	3.33	0.10	12	3.26	0.09	14	3.33	0.13	14	3.33	0.13	14	3.33	0.13
zrno	10	3.03	0.09	8	2.87	0.04	10	3.03	0.09	10	3.03	0.09	10	3.03	0.09
zvono	20	3.81	0.12	14	3.49	0.08	18	3.71	0.11	16	3.54	0.12	17	3.66	0.11

Appendix 3

Table of a) the number of clusters of related senses (N1) and number of senses (N2) listed in the dictionary *Rečnik Matice srpske*; the number of dictionary senses (N), entropy (H) and redundancy (T) of the sense probability distribution, for the dictionary senses; b) listed by the participants, c) with a mean familiarity rating of 2 or higher, d) with a mode familiarity rating above 1, and e) with a median familiarity rating above 1.

	a) Senses listed in the dictionary		b) Senses listed by participants			c) Senses with mean familiarity rating>1.9			d) Senses with mode familiarity rating >1			e) Senses with median familiarity rating >1		
	N1	N2	N	H	T	N	H	T	N	H	T	N	H	T
album	1	3	1	0.00	1.00	3	0.00	1.00	3	0.00	1.00	3	0.00	1
balkon	2	2	2	0.96	0.04	2	0.96	0.04	2	0.96	0.04	2	0.96	0.04
balon	4	6	4	1.90	0.05	5	1.90	0.18	4	1.90	0.05	5	1.90	0.18
berba	3	7	5	1.55	0.33	7	1.55	0.45	4	1.36	0.32	7	1.55	0.45
blok	7	7	5	2.18	0.06	6	2.18	0.16	5	2.18	0.06	6	2.18	0.16
boks	2	3	2	0.89	0.11	2	0.89	0.11	2	0.89	0.11	2	0.89	0.11
brada	4	5	4	1.64	0.18	5	1.64	0.29	3	1.31	0.17	5	1.64	0.29
ćelija	4	7	5	1.71	0.26	6	1.59	0.39	5	1.59	0.32	5	1.59	0.32
centar	8	9	7	2.40	0.15	9	2.40	0.24	8	2.40	0.20	9	2.40	0.24
ciklus	2	2	2	0.94	0.06	2	0.94	0.06	2	0.94	0.06	2	0.94	0.06
članak	5	7	5	1.87	0.19	6	1.87	0.28	4	1.25	0.38	6	1.87	0.28
crevo	3	3	2	1.00	0.00	3	1.00	0.37	2	1.00	0.00	3	1.00	0.37
čvor	8	11	7	2.35	0.16	10	2.35	0.29	9	2.35	0.26	10	2.35	0.29
đavo	4	12	8	2.36	0.21	12	2.36	0.34	12	2.36	0.34	12	2.36	0.34

	a) Senses listed in the dictionary		b) Senses listed by participants			c) Senses with mean familiarity rating>1.9			d) Senses with mode familiarity rating >1			e) Senses with median familiarity rating >1		
	N1	N2	N	H	T	N	H	T	N	H	T	N	H	T
dinar	2	2	2	0.99	0.01	2	0.99	0.01	2	0.99	0.01	2	0.99	0.01
disk	2	7	3	1.42	0.10	6	1.42	0.45	5	1.42	0.39	5	1.42	0.39
doktor	2	2	2	1.00	0.00	2	1.00	0.00	2	1.00	0.00	2	1.00	0
dugme	3	3	2	0.99	0.01	2	0.99	0.01	2	0.99	0.01	2	0.99	0.01
dvojka	6	9	2	0.89	0.11	5	0.89	0.62	3	0.89	0.44	5	0.89	0.62
figura	7	10	7	2.49	0.11	10	2.49	0.25	9	2.49	0.22	9	2.49	0.22
forma	6	6	5	1.95	0.16	6	1.95	0.24	6	1.95	0.24	6	1.95	0.24
gluma	3	4	3	1.58	0.00	4	1.58	0.21	4	1.58	0.21	4	1.58	0.21
govor	4	9	7	2.29	0.19	9	2.29	0.28	9	2.29	0.28	9	2.29	0.28
grada	6	7	5	1.90	0.18	7	1.90	0.32	6	1.90	0.27	7	1.90	0.32
grudi	4	4	2	1.00	0.00	4	1.00	0.50	4	1.00	0.50	4	1.00	0.5
grupa	3	5	4	1.80	0.10	5	1.80	0.23	5	1.80	0.23	5	1.80	0.23
guma	3	3	3	1.58	0.00	3	1.58	0.00	3	1.58	0.00	3	1.58	0
igrač	4	5	2	1.00	0.00	4	1.00	0.50	3	1.00	0.37	4	1.00	0.5
izbor	4	4	4	1.75	0.12	4	1.75	0.12	4	1.75	0.12	4	1.75	0.12
izlet	2	3	2	0.49	0.51	3	0.49	0.69	2	0.00	1.00	3	0.49	0.69
izraz	4	4	4	1.91	0.05	4	1.91	0.05	4	1.91	0.05	4	1.91	0.05
izvor	4	6	4	1.77	0.12	6	1.77	0.32	6	1.77	0.32	6	1.77	0.32
jezik	6	13	6	1.88	0.27	10	1.68	0.49	8	1.68	0.44	10	1.68	0.49

	a) Senses listed in the dictionary		b) Senses listed by participants			c) Senses with mean familiarity rating>1.9			d) Senses with mode familiarity rating >1			e) Senses with median familiarity rating >1		
	N1	N2	N	H	T	N	H	T	N	H	T	N	H	T
kanal	3	4	4	1.83	0.09	4	1.83	0.09	4	1.83	0.09	4	1.83	0.09
kapak	3	6	3	1.55	0.02	3	1.55	0.02	3	1.55	0.02	3	1.55	0.02
ključ	2	18	9	2.81	0.11	12	2.81	0.21	10	2.65	0.20	12	2.81	0.21
kljun	6	7	3	1.51	0.05	7	1.51	0.46	3	0.94	0.41	7	1.51	0.46
klub	1	2	2	1.00	0.00	2	1.00	0.00	2	1.00	0.00	2	1.00	0
klupa	5	6	3	1.16	0.27	3	0.99	0.37	3	0.99	0.37	3	0.99	0.37
kolač	6	9	3	1.08	0.32	5	1.08	0.54	4	1.08	0.46	4	1.08	0.46
komad	3	5	5	2.09	0.10	5	2.09	0.10	5	2.09	0.10	5	2.09	0.1
komora	5	9	5	1.97	0.15	7	1.84	0.34	4	1.51	0.24	8	1.84	0.39
korak	5	5	4	1.77	0.12	5	1.77	0.24	5	1.77	0.24	5	1.77	0.24
koren	6	9	7	2.54	0.10	9	2.54	0.20	8	2.54	0.15	9	2.54	0.2
korica	4	4	2	0.99	0.01	3	0.99	0.37	3	0.99	0.37	3	0.99	0.37
koža	3	4	4	1.45	0.27	4	1.45	0.27	3	1.24	0.22	4	1.45	0.27
krug	8	12	7	2.21	0.21	12	2.21	0.38	11	2.21	0.36	11	2.21	0.36
kruna	9	13	9	2.62	0.17	12	2.62	0.27	11	2.62	0.24	12	2.62	0.27
krzno	4	4	4	1.76	0.12	4	1.76	0.12	4	1.76	0.12	4	1.76	0.12
kurs	3	5	4	1.99	0.00	5	1.99	0.14	4	1.58	0.21	5	1.99	0.14
lanac	4	8	4	1.49	0.25	6	1.49	0.42	5	1.49	0.36	6	1.49	0.42
linija	9	13	10	2.79	0.16	12	2.70	0.25	12	2.70	0.25	13	2.79	0.25

	a) Senses listed in the dictionary		b) Senses listed by participants			c) Senses with mean familiarity rating>1.9			d) Senses with mode familiarity rating >1			e) Senses with median familiarity rating >1		
	N1	N2	N	H	T	N	H	T	N	H	T	N	H	T
list	8	10	5	2.05	0.12	10	2.05	0.38	8	2.05	0.32	9	2.05	0.35
lopta	4	5	4	1.87	0.06	5	1.87	0.19	5	1.87	0.19	5	1.87	0.19
loza	3	6	5	1.67	0.28	5	1.46	0.37	4	1.15	0.42	5	1.46	0.37
marka	4	5	3	1.56	0.02	3	1.56	0.02	3	1.56	0.02	3	1.56	0.02
masa	5	10	7	2.44	0.13	9	2.44	0.23	8	2.44	0.19	9	2.44	0.23
matica	9	11	8	2.77	0.08	11	2.77	0.20	10	2.77	0.17	10	2.77	0.17
mehur	4	7	6	2.15	0.17	7	2.15	0.23	6	1.81	0.30	7	2.15	0.23
mera	6	11	7	2.49	0.11	11	2.49	0.28	10	2.49	0.25	11	2.49	0.28
metar	6	6	3	1.17	0.26	6	1.17	0.55	5	1.17	0.50	6	1.17	0.55
minut	2	3	3	1.50	0.06	3	1.50	0.06	3	1.50	0.06	3	1.50	0.06
miris	3	3	3	1.47	0.07	3	1.47	0.07	3	1.47	0.07	3	1.47	0.07
model	5	8	6	2.34	0.09	8	2.34	0.22	8	2.34	0.22	8	2.34	0.22
momak	4	7	4	1.71	0.14	7	1.71	0.39	7	1.71	0.39	7	1.71	0.39
most	3	6	4	1.66	0.17	6	1.66	0.36	4	1.66	0.17	6	1.66	0.36
motiv	2	3	3	1.25	0.21	3	1.25	0.21	3	1.25	0.21	3	1.25	0.21
mreža	4	9	8	2.77	0.08	9	2.77	0.13	9	2.77	0.13	9	2.77	0.13
obim	2	2	2	1.00	0.00	2	1.00	0.00	2	1.00	0.00	2	1.00	0
oblast	5	7	3	1.57	0.01	6	1.57	0.39	6	1.57	0.39	6	1.57	0.39
oblik	9	11	8	2.31	0.23	11	2.31	0.33	10	2.31	0.31	10	2.31	0.31

	a) Senses listed in the dictionary		b) Senses listed by participants			c) Senses with mean familiarity rating>1.9			d) Senses with mode familiarity rating >1			e) Senses with median familiarity rating >1		
	N1	N2	N	H	T	N	H	T	N	H	T	N	H	T
obrada	4	4	4	1.75	0.13	4	1.75	0.13	4	1.75	0.13	4	1.75	0.13
obruč	6	8	5	2.19	0.06	8	2.19	0.27	8	2.19	0.27	8	2.19	0.27
ocena	0	4	3	1.55	0.02	4	1.55	0.22	4	1.55	0.22	4	1.55	0.22
oklop	4	4	3	1.54	0.03	4	1.54	0.23	4	1.54	0.23	4	1.54	0.23
organ	5	6	2	1.00	0.00	6	1.00	0.61	4	1.00	0.50	5	1.00	0.57
osnova	6	9	6	2.12	0.18	9	2.12	0.33	8	2.00	0.33	9	2.12	0.33
patent	4	6	4	1.54	0.23	6	1.54	0.40	5	1.54	0.34	6	1.54	0.4
pesak	2	2	2	0.63	0.37	2	0.63	0.37	2	0.63	0.37	2	0.63	0.37
pešak	3	3	3	1.39	0.12	3	1.39	0.12	3	1.39	0.12	3	1.39	0.12
pisak	4	10	5	2.19	0.06	7	2.19	0.22	5	2.19	0.06	7	2.19	0.22
pismo	4	8	4	1.69	0.16	8	1.69	0.44	6	1.53	0.41	8	1.69	0.44
platno	5	5	3	1.16	0.27	4	1.16	0.42	4	1.16	0.42	4	1.16	0.42
ploča	7	11	7	2.16	0.23	10	2.16	0.35	8	1.91	0.36	10	2.16	0.35
plod	3	3	3	1.57	0.01	3	1.57	0.01	3	1.57	0.01	3	1.57	0.01
pojas	9	14	6	2.30	0.11	13	2.30	0.38	10	2.30	0.31	13	2.30	0.38
pojava	6	7	4	1.86	0.07	7	1.86	0.34	7	1.86	0.34	7	1.86	0.34
pokret	4	7	5	2.26	0.03	7	2.26	0.20	7	2.26	0.20	7	2.26	0.2
polet	3	4	3	1.47	0.07	4	1.47	0.27	4	1.47	0.27	4	1.47	0.27
posao	4	9	5	1.91	0.18	9	1.91	0.40	9	1.91	0.40	9	1.91	0.4

	a) Senses listed in the dictionary		b) Senses listed by participants			c) Senses with mean familiarity rating>1.9			d) Senses with mode familiarity rating >1			e) Senses with median familiarity rating >1		
	N1	N2	N	H	T	N	H	T	N	H	T	N	H	T
poskok	2	2	2	1.00	0.00	2	1.00	0.00	2	1.00	0.00	2	1.00	0
pošta	5	8	4	1.80	0.10	7	1.80	0.36	4	1.80	0.10	6	1.80	0.3
potez	6	9	5	2.09	0.10	9	2.09	0.34	7	2.09	0.26	9	2.09	0.34
poziv	3	6	6	2.39	0.08	6	2.39	0.08	5	2.11	0.09	6	2.39	0.08
prašak	5	6	3	1.06	0.33	6	1.06	0.59	2	0.86	0.14	5	0.86	0.63
pravac	4	7	5	2.08	0.10	6	2.08	0.20	6	2.08	0.20	6	2.08	0.2
prenos	4	5	4	1.72	0.14	5	1.72	0.26	5	1.72	0.26	5	1.72	0.26
prilog	4	4	4	1.71	0.15	4	1.71	0.15	4	1.71	0.15	4	1.71	0.15
profil	5	5	3	1.36	0.14	5	1.36	0.41	5	1.36	0.41	5	1.36	0.41
pruga	7	12	4	1.61	0.19	11	1.61	0.53	9	1.39	0.56	12	1.61	0.55
račun	7	13	9	2.79	0.12	13	2.79	0.25	12	2.63	0.27	13	2.79	0.25
radnja	3	15	7	2.32	0.17	14	2.32	0.39	12	2.32	0.35	14	2.32	0.39
rebro	7	8	3	1.26	0.20	6	1.26	0.51	5	1.26	0.46	6	1.26	0.51
salon	3	5	5	2.09	0.10	5	2.09	0.10	5	2.09	0.10	5	2.09	0.1
samica	3	4	3	1.27	0.20	4	1.27	0.36	4	1.27	0.36	4	1.27	0.36
sastav	6	6	6	2.20	0.15	6	2.20	0.15	5	2.10	0.09	5	2.10	0.09
savet	5	6	5	1.89	0.19	6	1.89	0.27	6	1.89	0.27	6	1.89	0.27
scena	3	6	6	2.22	0.14	6	2.22	0.14	6	2.22	0.14	6	2.22	0.14
servis	3	4	4	1.80	0.10	4	1.80	0.10	4	1.80	0.10	4	1.80	0.1

	a) Senses listed in the dictionary		b) Senses listed by participants			c) Senses with mean familiarity rating>1.9			d) Senses with mode familiarity rating >1			e) Senses with median familiarity rating >1		
	N1	N2	N	H	T	N	H	T	N	H	T	N	H	T
sfera	4	5	4	1.75	0.12	5	1.75	0.24	5	1.75	0.24	5	1.75	0.24
signal	2	3	2	0.85	0.15	3	0.85	0.47	3	0.85	0.47	3	0.85	0.47
sila	12	18	11	3.06	0.12	18	3.06	0.27	17	3.06	0.25	18	3.06	0.27
sistem	8	12	10	2.85	0.14	12	2.85	0.20	12	2.85	0.20	12	2.85	0.2
skakač	3	4	3	1.58	0.01	4	1.58	0.21	4	1.58	0.21	4	1.58	0.21
skok	4	9	4	1.77	0.11	9	1.77	0.44	9	1.77	0.44	9	1.77	0.44
škola	4	7	4	1.58	0.21	7	1.58	0.44	7	1.58	0.44	7	1.58	0.44
sloj	2	2	2	0.97	0.03	2	0.97	0.03	2	0.97	0.03	2	0.97	0.03
sluh	4	4	3	1.25	0.21	4	1.25	0.37	3	1.25	0.21	3	1.25	0.21
smer	4	6	4	1.76	0.12	6	1.76	0.32	4	1.76	0.12	6	1.76	0.32
snimak	2	3	2	0.95	0.05	3	0.95	0.40	3	0.95	0.40	3	0.95	0.4
stav	6	10	8	2.22	0.26	9	2.06	0.35	8	2.06	0.31	8	2.06	0.31
stena	4	6	3	1.49	0.06	5	1.49	0.36	4	1.49	0.25	5	1.49	0.36
stepen	8	14	6	2.29	0.11	14	2.29	0.40	12	2.29	0.36	14	2.29	0.4
stopa	9	14	6	2.25	0.13	13	2.25	0.39	7	2.14	0.24	13	2.25	0.39
struja	6	8	5	2.00	0.14	8	2.00	0.33	6	1.89	0.27	7	2.00	0.29
struk	3	10	5	1.92	0.17	7	1.72	0.39	5	1.72	0.26	7	1.72	0.39
tabla	3	7	4	1.76	0.12	5	1.76	0.24	4	1.76	0.12	5	1.76	0.24
tačka	8	11	9	2.85	0.10	10	2.85	0.14	10	2.85	0.14	10	2.85	0.14

	a) Senses listed in the dictionary		b) Senses listed by participants			c) Senses with mean familiarity rating>1.9			d) Senses with mode familiarity rating >1			e) Senses with median familiarity rating >1		
	N1	N2	N	H	T	N	H	T	N	H	T	N	H	T
talas	3	6	5	2.10	0.10	6	2.10	0.19	6	2.10	0.19	6	2.10	0.19
tekst	3	5	2	0.70	0.30	5	0.70	0.70	4	0.70	0.65	5	0.70	0.7
termin	3	3	2	1.00	0.00	3	1.00	0.37	2	1.00	0.00	3	1.00	0.37
teza	4	6	3	1.58	0.00	5	1.58	0.32	5	1.58	0.32	5	1.58	0.32
tezga	2	2	2	1.00	0.00	2	1.00	0.00	2	1.00	0.00	2	1.00	0
traka	5	5	4	1.73	0.13	5	1.73	0.25	3	1.23	0.23	5	1.73	0.25
ugao	4	6	3	1.53	0.04	6	1.53	0.41	6	1.53	0.41	6	1.53	0.41
ukus	2	4	4	1.93	0.03	4	1.93	0.03	4	1.93	0.03	4	1.93	0.03
uslov	2	2	2	0.90	0.10	2	0.90	0.10	2	0.90	0.10	2	0.90	0.1
uspon	4	6	3	1.32	0.17	6	1.32	0.49	6	1.32	0.49	6	1.32	0.49
vatra	9	11	6	2.25	0.13	11	2.25	0.35	9	1.97	0.38	11	2.25	0.35
venac	11	11	7	2.34	0.17	10	2.34	0.29	7	2.04	0.27	10	2.34	0.29
veza	5	11	9	2.86	0.10	11	2.86	0.17	11	2.86	0.17	11	2.86	0.17
visina	8	9	9	2.69	0.15	9	2.69	0.15	9	2.69	0.15	9	2.69	0.15
vrat	2	2	2	0.97	0.03	2	0.97	0.03	2	0.97	0.03	2	0.97	0.03
vreća	3	5	1	0.00	1.00	4	0.00	1.00	4	0.00	1.00	4	0.00	1
zakon	7	8	5	2.08	0.10	8	2.08	0.31	6	2.08	0.19	8	2.08	0.31
zglob	3	7	3	1.46	0.08	5	1.46	0.37	3	0.84	0.47	5	1.46	0.37
žica	8	11	6	2.43	0.06	10	2.43	0.27	6	2.43	0.06	9	2.43	0.23

NUMBER, RELATIVE FREQUENCY, ENTROPY, REDUNDANCY …

	a) Senses listed in the dictionary		b) Senses listed by participants			c) Senses with mean familiarity rating>1.9			d) Senses with mode familiarity rating >1			e) Senses with median familiarity rating >1		
	N1	N2	N	H	T	N	H	T	N	H	T	N	H	T
znak	7	8	5	2.02	0.13	8	2.02	0.33	8	2.02	0.33	8	2.02	0.33
zrno	3	4	3	1.53	0.03	4	1.53	0.23	4	1.53	0.23	4	1.53	0.23
zvono	2	4	3	1.55	0.02	3	1.55	0.02	3	1.55	0.02	3	1.55	0.02