PRIMENJENA PSIHOLOGIJA, 2019, Vol. *12*(3), STR. 283-304

UDK: 159.9.07:811.163.41
Originalni naučni rad
doi: 10.19090/pp.2019.3.283-304

**Dušica Filipović Đurđević**[1]

Department of Psychology, Faculty of Philosophy, University of Belgrade

Laboratory for Experimental Psychology, Faculty of Philosophy, University of Belgrade

Laboratory for Experimental Psychology, Faculty of Philosophy, University of Novi Sad

[1] Corresponding author email: dusica.djurdjevic@f.bg.ac.rs

Primljeno: 20. 06. 2019.
Primljena korekcija: 01. 09. 2019.
Prihvaćeno za štampu: 15. 09. 2019.

# BALANCE OF MEANING PROBABILITIES IN PROCESSING OF SERBIAN HOMONYMY[2]

The research deals with the set of Serbian homonymous nouns (nouns with multiple unrelated meanings) presented in the norming study and in the visual lexical decision task experiment. Native speakers listed the meanings of homonymous words and provided word familiarity and word concreteness ratings. Accordingly, the first database of Serbian homonyms was constructed containing subjective meanings of homonymous nouns along with the estimated meaning probabilities, as well as a number of meanings, redundancy and entropy of the distribution of meaning probabilities, word familiarity and word concreteness. The processing disadvantage of homonymous nouns over unambiguous nouns was replicated in the visual lexical decision task. Additionally, the processing of homonymous nouns was linked with redundancy: the information theory measure of the balance of meaning probabilities. The results revealed that homonyms with higher redundancy of the meaning probability distribution (i.e., unbalanced meaning probabilities) were processed faster. This finding was in accordance with the hypothesis derived from the Semantic Settling Dynamics account of the processing of ambiguous words, according to which the competition among the unrelated meanings derived the processing disadvantage in homonymy. However, the same pattern was not observed for the number of meanings and entropy, inviting for further research of the processing of ambiguous words.

**Keywords**: entropy, homonymy, number of meanings, redundancy, visual lexical decision task

[2] This work was supported by Ministry of Education, Science and Technological Development of the Republic of Serbia (grant number: 179033 and 179006).

## Introduction

Processing of lexical ambiguity has long been the subject of psycholinguistic investigations. Early work was dominated by equivocal findings: the researchers observed both processing advantage and processing disadvantage of ambiguous words, and sometimes the lack of ambiguity effect has been reported (Azuma & Van Orden, 1997; Borowsky & Masson, 1996; Clark, 1973; Gernsbacher, 1984; Hino & Lupker, 1996, 2002; Jastrzembski, 1981; Kellas, Ferraro, & Simpson, 1988; Millis & Bution, 1989; Rubenstein, Garfield, & Millikan, 1970). The first step in resolving this inconsistency was accomplished by Rodd, Gaskell, and Marslen-Wilson (2002), who demonstrated that the type of lexical ambiguity was important for processing effects. They pointed to the difference between homonyms, words with unrelated meanings (e.g., *river bank* and *financial bank*), and polysemes, words with multiple related senses (e.g., *daily paper* and *paper as a material*). Compared to unambiguous words (words with one meaning and one sense), polysemes were processed faster, whereas homonyms took more time to be recognized in a lexical decision task.

The observed asymmetry in processing effects of polysemy and homonymy was accounted for by parallel distributed model proposed by Rodd, Gaskell, and Marslen-Wilson (2004), and a similar upgraded model named Semantic Settling Dynamics (SSD) model proposed by Armstrong and Plaut (2016; and described in more detail by Armstrong, 2012). The model learnt to map the form to the meaning of unambiguous words (one-to-one mapping), homonymous words (one form to many unrelated meanings, which did not share units at the semantic level), and polysemous words (one form to many related senses, which shared units at the semantic level). In the model, the delay in recognition time (i.e., a greater number of simulation cycles) for the homonyms relative to the unambiguous words was attributed to the competition among the unrelated meanings at the semantic level (due to inhibitory connections among the units which represented those meanings). On the other hand, faster recognition of polysemous words compared to the unambiguous words was attributed to the wider distribution of representational units (compared to one-to-one mapping of unambiguous words) and the reduced competition among them (as the related senses shared features, i.e., share units at the semantic level). However, in the case of polysem, competition would arise later in the processing, as also predicted by the model (Armstrong, 2012; Armstrong & Plaut, 2016; Rodd et al., 2004), and observed in the empirical data (for a review see Edington & Tokowicz, 2015, and also Armstrong, 2012; Armstrong & Plaut, 2016).

The research conducted in Serbian language has confirmed processing advantage of polysemous words in the visual lexical decision task as compared to the processing of unambiguous words, as well as the facilitation effect of the number of senses (Filipović Đurđević & Kostić, 2008). However, the research in Serbian language has also demonstrated the predictive power of another variable, a balance of the sense frequencies (Filipović Đurđević, 2007; Filipović Đurđević & Kostić, under review). In this research, polysemy was defined as the sense uncertainty, and described numerically in terms of information theory measure of entropy, as previously suggested by Gilhooly and Loggie (1980). Entropy (*H*) of the word *w* is the measure of an average uncertainty within a set of word senses (Cover & Thomas, 1991; Filipović Đurđević & Kostić, 2017; Shannon, 1948):

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

In (1) $p_i$ denotes the probability of the word sense (i.e., the relative frequency of usage), and *n* denotes the number of senses of the given word *w*. Entropy of the distribution of probabilities (relative frequencies) of the word senses is positively correlated to the number of word senses: the higher the number of senses, the higher the entropy (i.e., uncertainty of the word's true senses). However, in addition to the number of senses, the uncertainty is also influenced by the relative frequencies of the senses, i.e., by the balance of sense probabilities. Information theory measure that describes exactly this aspect of uncertainty is called redundancy (Cover & Thomas, 1991; Filipović Đurđević & Kostić, 2017; Shannon, 1948):

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

In (2) *T(w)* denotes the redundancy of the word *w*, *H(w)* denotes the entropy of the word *w*, and *N* stands for the number of senses of the given word. The higher the balance among the sense probabilities (i.e., the smaller the differences among them), the smaller the redundancy should be, and vice versa.

As described in Filipović Đuđević and Kostić (2017, and also Filipović Đurđević, 2007; Filipović Đurđević & Kostić, under review), it can be demonstrated that entropy can be reduced to the number of senses and redundancy, and that the two can be used interchangeably. To summarize:the larger the number of senses and the more balanced the sense frequencies (smaller the redundancy), the larger the sense uncertainty (higher the entropy). Experiments with presenting polysemous nouns in visual lexical decision task confirmed that higher uncertainty led to shorter processing latencies, and as predicted (based on the presented equations), that larger number of senses and smaller redundancy led to faster processing.

The findings of Filipović Đurđević (2007) and Filipović Đurđević and Kostić (under review) fit with the SSD lexical ambiguity account described by Armstrong (2012) and Armstrong and Plaut (2016). According to this account, early processing which was captured by visual lexical decision task was characterized by the strong polysemy advantage over the unambiguous words. In their model, the authors presented the case of the "ideally ambiguous words", as they called it. Those would be the polysemous words with a large number of senses of equal probability. The imbalance among the sense probabilities would make that word less ambiguous compared to the word with the same number of perfectly balanced sense frequencies, i.e., it would make it more similar to the unambiguous words. Therefore, according to the model, the processing time of that word should also be more similar to the processing time of the unambiguous word. Hence, the model would predict slower processing for the words with unbalanced sense frequencies, as observed in the experiments (Filipović Đurđević, 2007; Filipović Đurđević & Kostić, under review). The same holds for the number of senses effect (a word with fewer senses is more similar to the unambiguous word), as well as the effect of the entropy, as the overall measure of the sense uncertainty.

The question remained whether the uncertainty as described by the information theory measures affected also the processing time of homonymous words, and if it did, in what direction. Unlike polysemy, where previous research did not look into the relation of the balance of the sense probabilities and the processing time of isolated words, there were reports on the balance of probabilities effects of the homonymous words meanings on the processing time. For example, Armstrong, Tokowicz, and Plaut (2012) operationalized the level of balance of meaning frequencies by a measure named β, which presented the difference between the probability of the dominant (most frequent) meaning, and the second most frequent meaning. They demonstrated that this measure was a significant predictor of the processing latencies and accuracy in recognizing homonyms in a visual lexical decision task. Unlike β, the measure proposed by Gilhooly and Logie (1980), labeled as U, which was calculated as entropy of the word meanings, was not significant in predicting processing of homonyms. However, to the best of our knowledge, processing of homonyms was not linked to redundancy, as an information theory measure of the balance of meaning probabilities. Larger values of β should be linked to larger values of redundancy. In Armstrong et al. (2012) almost all variations in meaning probabilities were accounted for by the difference between the dominant and the second meaning. Therefore, in a different set of homonyms, redundancy should not only match, but also outperform the predictive power of β, because it captured the variations in meaning probabilities across the full set, and not only for the two most dominant meanings. According to the SSD account (Armstrong, 2012; Armstrong & Plaut, 2016; but also Rodd et al., 2004) words with a lower number of unrelated meanings, and words with less balanced frequencies of unrelated meanings, i.e., words that were less homonymous (hence more similar to the unambiguous words) should elicit processing latencies that were also more similar to the processing time of unambiguous words compared to ideally homonymous words depicted by the model (words with a large number of perfectly balanced unrelated meanings). Therefore, an increase in entropy should be followed by the increase in processing time. Along the same line, the increase in a number of meanings should be followed by the increase in the processing time, whereas the increase in redundancy should be followed by the decrease in the processing time.

## The Present Study

The first aim of this study was to replicate the homonymy effect, i.e., the slower processing of homonyms as compared to the processing of unambiguous words in Serbian language. The next aim was to test the predictive power of information theory measures in the processing of homonymy. Based on the SSD account of Armstrong and Plaut (2016) it could be predicted that the sense uncertainty (operationalized as entropy in this study) should exhibit reversed effect to that observed in polysemy, i.e., it should be positively correlated with processing latencies in the visual lexical decision task. Along the same line, it can be predicted that the number of meanings should be positively correlated with reaction time in this task (the larger the number of meanings, the slower the recognition of the word), and that redundancy should be negatively correlated with reaction time (the less balanced the meanings, the faster the processing).

In order to accomplish the set goals, a norming study was conducted to estimate the values of the predictors in question (a number of meanings, redundancy, and entropy, along with relevant control variables), as well as a visual lexical decision task experiment to collect processing latencies for a set of Serbian homonyms and unambiguous words.

## Norming Study

In the first phase of the study, stimuli (homonymous and unambiguous words) were selected from the dictionary, while the ratings on word familiarity (Balota, Pilotti, & Cortese, 2001; Gernsbacher, 1984) and word concreteness (Paivio, Yuille, & Madigan, 1968) were collected for the selected items. For the homonyms, word meanings were collected from the native speakers. The meaning probabilities were estimated based on the listed meanings (as the proportion of participants listing the given meaning relative to the total number of participants). Finally, the number of meanings (by counting the categories of meanings listed by participants), redundancy (based on equation 2), and entropy (based on equation 1) were calculated based on these probabilities.

## Method

**Participants.** Familiarity ratings were collected from a group of 20 native speakers of Serbian, and concreteness ratings were collected from another group of 20 participants. The third group of 72 Serbian native speakers, who were randomly assigned to one of the two word lists, and one of the three random word orders within a list, provided meanings for the homonymous words. All of them participated for the partial course credit or as volunteers from the pool of students at the Department of Psychology, Faculty of Philosophy, University of Novi

Sad, and the Department of Psychology, Faculty of Philosophy, University of Belgrade. All participants signed the information consent forms for the study that was approved by the Ethical Committee of the Department of Psychology at the Faculty of Philosophy, University of Novi Sad.

**Materials and Design.** Stimuli were selected based on the number of entries in Rečnik Matice Srpske dictionary of Serbian language (RMS, 1967-1976). Following Rodd et al. (2002) words with multiple entries were considered as homonyms, and words with a single entry were considered as non-homonyms. In order to exclude strictly polysemous words from the study, out of non-homonyms, only words with a single listed sense were considered. Following this criterion, 46 homonyms and 46 unambiguous nouns were selected. Within the set of homonyms, some entries had single, and some entries had multiple listed senses. We used this information to split the homonyms into the subset of strictly homonymous nouns (with one sense per entry), and the subset of hybrid nouns, homonymy with polysemous meanings (with multiple senses per entry).

**Procedure.** We describe separately the three procedures.

***Word familiarity (subjective frequency) ratings.*** Participants were presented a list of 92 words with a seven-point scale printed next to each word. They were instructed to rate to what extent they were familiar with the word, i.e., how often they encountered it. The scale ranged from a completely unfamiliar word (i.e., the word that they never encountered; 1) to very familiar word (i.e., the word that they encountered very frequently; 7). In order to control the serial effects, there were three parallel versions of the same list, according to three different random orders of the words. Each participant was presented only one random order. It took approximately 10 minutes to complete the booklet.

***Word concreteness ratings.*** The procedure here was the same as with word familiarity, except that the participants were instructed to rate to what extent they could perceptually experience (see, hear, touch, smell, or taste) an object denoted by a given word. The scale ranged from abstract (impossible to perceptually experience; 1) to concrete (highly possible to perceptually experience; 7).

***Word meaning elicitation.*** The list of 46 homonyms was divided into two lists. Every sub-list was printed in a separate booklet, in three parallel forms, according to three random orders, to control the serial effects. The participants were presented with a detailed instruction to list all of the meanings of homonymous words they could think of (total meaning metric; Azuma, 1996), followed by an example with a word (that was not included in the list), and its meanings from the dictionary. They were encouraged to use definitions, sentence examples, along with all means available to describe the meaning they had in mind. Participants were randomly assigned to one of these lists and it took them approximately 20 minutes to finish the task.

## Results and Discussion

Descriptive statistics for the variables addressed in the norming study are presented in Table 1. Entropy, a number of meanings and redundancy were estimated in two ways: by counting all of the meanings listed by participants, and by counting only the meanings that were listed by at least two participants (to avoid idiosyncratic meanings).

The full list of words, along with the collected ratings, and all the relevant measures are available in the Appendix (Table A). The inspection of dispersion measures revealed that the sample size was reasonably estimated. Additionally, the database containing individual meanings of words and their frequencies can be downloaded from this link: https://osf.io/nxw23/?view_only=6e93c417a0c4 4b3fb5af315ad97347f2.

Table 1
*Descriptive measures of the homonymous nouns from the database*

| | *M* | *SD* | *Min* | *Max* |
|---|---|---|---|---|
| Lemma frequency (in two millions) | 100.83 | 233.04 | 0.10 | 1452.00 |
| Word length in letters | 3.89 | 0.97 | 3.00 | 6.00 |
| Coltheart's N | 9.33 | 6.97 | 0.00 | 28.00 |
| Word familiarity | 5.74 | 0.90 | 3.60 | 6.84 |
| Word concreteness | 5.08 | 1.49 | 2.05 | 6.85 |
| Entropy (raw) | 0.34 | 0.11 | 0.12 | 0.58 |
| Number of meanings (raw) | 2.98 | 1.16 | 2.00 | 6.00 |
| Redundancy (raw) | 0.22 | 0.17 | 0.00 | 0.60 |
| Entropy | 0.33 | 0.10 | 0.12 | 0.58 |
| Number of meanings | 2.63 | 0.74 | 2.00 | 5.00 |
| Redundancy | 0.18 | 0.17 | 0.00 | 0.60 |

*Note. M* - mean; *SD* – standard deviation.

# Experiment

## Method

**Participants.** Forty-six students from the Faculty of Philosophy, University of Novi Sad, and the Faculty of Philosophy, University of Belgrade took part in this experiment for a partial course credit. All of them were native speakers of Serbian and all of them had normal or corrected-to-normal vision. All participants signed the information consent form prior to participating in the study which was approved by the Ethical Committee of the Department of Psychology at the Faculty of Philosophy, University of Novi Sad.

**Materials and Design.** Based on the collected norms, 37 homonyms and 37 unambiguous words were selected. The two groups were matched for the word length in letters, lemma frequency (Kostić, 1999), familiarity (subjective frequency), concreteness, and orthographic neighbourhood size (Coltheart's *N*; Coltheart, Davelaar, Jonasson, & Besner, 1977). The matching procedure was based on the pair-wise matching as much as possible, and was finalized by the matching of group means. Statistical testing (t-test) revealed no significant differences between the two groups of words. Additional 9 homonyms from the norming study were presented (and analysed in the second part of the analysis), as well as 9 additional filler nouns. The additional homonyms were not included in the first set as we were not able to find the matching unambiguous words. However, they were presented in the experiment to increase the power of the second part of the analysis). Finally, a list of 92 pseudo-words was constructed by a native speaker of Serbian to mimic the orthographical/phonological structure of the selected words.

**Procedure.** Participants were engaged in a visual lexical decision task, responding by mouse button press. Stimuli were presented on a computer screen by using OpenSesame experimental software (Mathôt, Schreij, & Theeuwes, 2012). Each stimulus was preceded by a fixation point for 1500 ms and remained on the screen until the participant responded or until the end of 1500 ms interval. Each participant was presented a different, random order of stimuli. Prior to the experiment, participants were presented 20 practice trials that were not a part of the experimental stimuli list and were not included in the analysis.

## Results and Discussion

Prior to the analysis, data from participants and items with below 20% of accuracy were excluded. A reaction time was inversely transformed as suggested by Baayen and Milin (2010), and all numerical predictors were standardized (transformed to z-scores), as suggested by Gelman and Hill (2007). The data were analysed in R statistical software (R core team, 2019), by applying Generalized Additive Mixed Models, using *mgcv* (Wood, 2006, 2011) and *itsadug* (van Rij, Wieling, Baayen, & van Rijn, 2016) packages. The plots were produced in the package *ggplot2* (Wickham, 2016). In the analysis of numerical predictors, the importance of each predictor was tested by a model comparison. The final model, which included only the predictors justified by the data, was refitted by excluding the influential data points. The two versions of the model did not differ, and the refitted model was reported. In addition to the fixed effects, we also included the random effects in the model. We fitted random intercepts for the items, and the factor smooth for the order of trial presentation for every participant. By doing so, we took into account the overall speed of every participant along with their individual oscillations in speed during the course of the experiment.

In the first step, we compared processing latencies of homonymous and unambiguous words that were matched on several control variables (as described in the Method section). Ambiguity was attested as a significant predictor of the reaction time, as reported in Table 2. As predicted, unambiguous words were recognized faster as compared to homonyms (Figure 1).

Table 2
*Coefficients from the generalized additive model of ambiguity fitted to response latencies*

| Parametric coefficients: | | | | |
|---|---|---|---|---|
| | Estimate | Std. Error | *t* | *Pr(>\|t\|)* |
| Intercept | -1.585 | .036 | -43.752 | <.0001 |
| Order of trial presentation | -.0002 | .0002 | -1.031 | .302 |
| Ambiguity: unambiguous words | -.068 | .033 | -2.074 | .038 |
| Smooth terms: | | | | |
| | edf | Ref.df | *F* | *p* |
| By-Participant factor smooth for Order of Trial Presentation | 84.680 | 341 | 2.845 | <.0001 |
| By-Item random intercept | 54.480 | 60 | 9.820 | <.0001 |

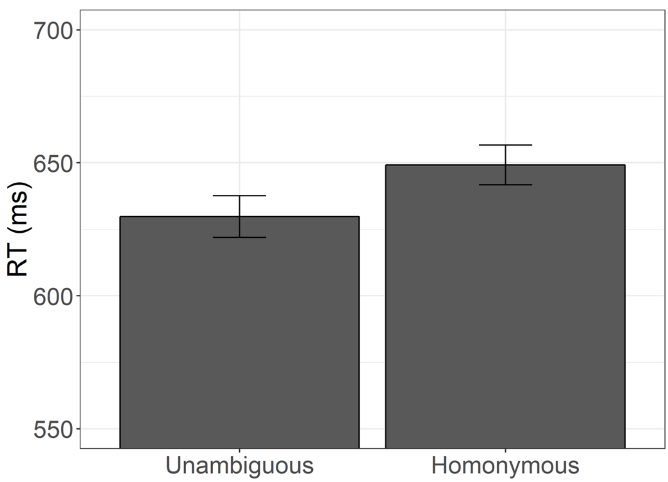


*Figure 1*. Processing latencies for the group of Unambiguous and Homonymous words, as observed in the experiment (vertical bars denote 95% of confidence intervals).

In the next step, we selected only homonymous words in order to investigate the effects of information theory based measures of lexical ambiguity. As described earlier, these words were split into two groups based on the finding that some of them elicited only homonymous meanings, whereas some of them elicited polysemous senses embedded in some of their homonymous meanings. This variable was named 'Type of ambiguity', and was also included in the analyses. In addition to control variables, the first model included the meaning entropy. However, the effect of this predictor was not significant. In the second model, entropy was replaced with the number of meanings and redundancy. The analysis revealed that only redundancy was predictive of processing latencies, as presented in Table 3. We observed a significant interaction of the type of ambiguity, word familiarity (subjective frequency) and redundancy, which revealed that the effect of redundancy was modulated by familiarity differently for two types of ambiguity (as depicted in Figure 2). For words with strictly homonymous meanings (Figure 2, left panel) there was a strong facilitation effect of redundancy that was most prominent for words of higher familiarity, and weakened with a decrease in familiarity. On the other hand, the effect of redundancy was not present for words with a mixture of homonymy and polysemy (Figure 2, right panel).

Table 3
*Coefficients from the generalized additive model of numerical predictors fitted to response latencies to homonyms*

| Parametric coefficients: | | | | |
|---|---|---|---|---|
| | Estimate | Std. Error | *t* | *Pr(>\|t\|)* |
| Intercept | -1.588 | .027 | -58.541 | <.0001 |
| Order of trial presentation | -.004 | .008 | -.502 | .616 |
| **Smooth terms:** | | | | |
| | edf | Ref.df | *F* | *p* |
| Tensor product smooth for Redundancy and Word Familiarity (subjective frequency) at the level of strictly homonymous words | 4.259 | 4.438 | 6.872 | <.0001 |
| Tensor product smooth for Redundancy and Word Familiarity (subjective frequency) at the level of homonyms with polysemous meanings | 3.000 | 3.000 | 16.781 | <.0001 |
| By-Participant factor smooth for Order of Trial Presentation | 56.072 | 341 | 1.499 | <.0001 |
| By-Item random intercept | 19.681 | 29 | 2.23 | <.0001 |

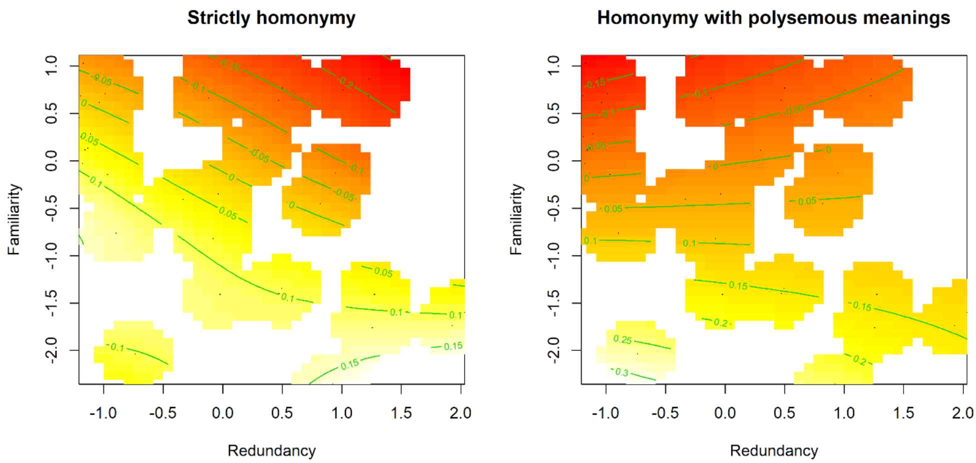


*Figure 2*. The interaction of ambiguity type, word familiarity (subjective frequency), and redundancy of the distribution of meaning probabilities (higher redundancy, less balanced probabilities). The panel on the left hand side plots familiarity by redundancy interaction for words with strictly homonymous meanings, and the panel on the right hand side plots familiarity by redundancy interaction for homonyms with polysemy within certain meanings. Reaction latencies are colour coded: brighter shades represent slower processing, whereas darker shades represent shorter latencies. The lines plot areas with the same values of the reaction time (its negative inverse).

## General Discussion

This research presented a norming study for the set of Serbian homonyms and a visual lexical decision task bringing two important insights. It replicated processing disadvantage of homonymous words in a novel language and revealed that redundancy of the distribution of unrelated word meanings predicted recognition time of an isolated homonym.

Firstly, a database of the norms for familiarity, concreteness, a number of meanings of 46 Serbian homonymous nouns was presented, along with the associated meaning probabilities and information theory measures of entropy and redundancy that were derived based on those probabilities. Although the number of homonyms in the database was not very large, this database was the first of its kind for Serbian language (in addition to the similar database for Serbian polysemous nouns that was collected by Filipović Đurđević & Kostić, 2017). It was constructed in a manner of similar databases that were constructed for other languages (Armstrong et al., 2012; Azuma, 1996; Ferraro & Kellas, 1990; Gawlick-Grendell & Woltz, 1994; Rice, Beekhuizen, Dubrovsky, Stevenson, & Armstrong, 2018; Twilley, Dixon, Taylor, & Clark, 1994), and enriched by the estimate of the distribution redundancy of the meaning probabilities. On the one hand, the size of the database was affected by the restriction of using strictly homonymous nouns, i.e., homonyms that represented nouns in all their meanings. This criterion was not applied in other databases, as they sometimes included words that belonged to different part of speech categories depending on the meaning. On the other hand, the size of the database was restricted due to the nature of language, namely the fact that homonyms were rare in language (Rodd et al., 2002). This is a consequence of the way homonyms are being developed:  the accidental overlap in a form between the word borrowed from the foreign language and the word that is already present in the given language (e.g., *pop*). Therefore, in spite of its size, in addition to enabling the current study, the presented database could be a useful resource for future linguistic and psycholinguistic studies, especially in a language with not enough unabridged up-to-date dictionaries.

Next, the processing disadvantage of homonymous words, as observed in Rodd et al. (2002) is replicated in Serbian language. A group of homonymous words is presented along with a group of unambiguous controls to demonstrate that homonyms take more time than unambiguous words to be recognized in a visual lexical decision task. This finding is significant as the effect of homonymy has proven to be unstable and not an easy one to detect (Eddington & Tokowicz, 2015; Medeiros & Armstrong, 2017), as it adds to the body of the research that has demonstrated the same difference (Armstrong & Plaut, 2016; Rodd et al., 2002). Also, it fits with the prediction of the computational models of lexical ambiguity, such as the one proposed by Rodd et al. (2004), and the model that has been built upon it, SSD model of Armstrong and Plaut (2012). According to these models, the activation of semantic representations of unambiguous words steadily raises as the semantic processing unfolds over time. This increase in activation is fed by the accumulating evidence from the output that flows in the one-to-one mapping from the form to the meaning. The same activation for the homonymous words follows a similar (although not identical) raising trend, but it is of much lower intensity as compared to the activation of words with only one meaning. The reason for this diminished activation lies in the fact that the form-to-meaning mapping is of the one-to-many kind. Importantly, the many meanings at the semantic level are mutually unrelated, i.e., they represent separate units at the semantic level. Consequently, being linked by inhibitory connections, the feeding of activation inevitably leads to the competition among them. The time taken for this competition to be resolved is the added time needed for the homonym to be recognized as the familiar word, as observed in this study.

Finally, homonymous nouns were analysed in two parallel statistical models. The first model which included information theory measure of entropy revealed that this predictor did not account for variation in processing latencies over and above the control variables. The second model, which included the number of meanings and the redundancy of the distribution of meaning probabilities revealed only the significant effect of redundancy. Based on computational model of Rodd et al. (2004), as well as the SSD model of Armstrong and Plaut (2016) the prediction was derived that the increase in uncertainty of the true homonymous meaning would be followed by the increase in processing latencies in the visual lexical decision task. More precisely, the larger the number of meanings and the smaller the redundancy, and consequently the higher the entropy, the longer the processing should be. This prediction derives from the fact that the model depicts the behaviour of the "perfect homonyms", i.e., homonyms with many equally frequent meanings, and implies that homonyms with fewer meanings, as well as homonyms with unbalanced meaning frequencies, are less ambiguous, and thus more similar to unambiguous words. Given that the model predicts less activation (and thus slower processing) for "perfect homonyms" as compared to unambiguous words, the prediction that the increase in the level of ambiguity (larger number of meanings, lower redundancy, and consequently higher entropy) would be followed by slower recognition appears straightforwardly. However, that is not exactly what the results have revealed. A lack of the number of meanings effect could be attributed to the restriction in range. For example, the number of senses ranges from 2 to 18 (or even 35, depending on the method of estimation) in a similar study on polysemous words, whereas the number of meanings ranges from 2 to 4 in this study. The lack of redundancy (imbalance of meaning frequencies) effect for homonyms with polysemous senses nested within meanings could also be explained by opposite effects of redundancy of senses and redundancy of meanings which eliminate each other. However, the observed facilitation effect of redundancy of words with strictly homonymous meanings is completely in line with the prediction derived from the model. Armstrong et al. (2012) and Rice et al. (2018) described a similar effect of the imbalance measure of meaning frequencies similar to redundancy, the β, or "biggest", as they called it. In that study, the authors observed that homonyms with larger frequencies of the dominant meaning were processed faster. However, although their β and the redundancy were highly correlated, the two measures should not be identified, as redundancy captured the variability of frequencies across the full range of meanings. At the same time, the same authors reported the analysis that revealed the non-significant effect of entropy (U, as defined by Gilhooly & Logie, 1980). The entropy effect was expected to be observed as the inevitable consequence of the effects of a number of meanings and redundancy acting in the additive manner. The lack of entropy effect remains puzzling at the moment, especially in the light of the previous finding that the effect of the sense uncertainty measures on processing of polysemous words is in the perfect accordance with the model predictions (Filipović Đurđević & Kostić, under review). Nevertheless, the observed effect of redundancy of probability distribution of homonymous meanings should be considered as highly informative of the current models of lexical ambiguity processing and investigated further. A potential approach would be to investigate it from the point of view of the model based on discriminative learning and contextual diversity, as already applied to the polysemous words (Filipović Đurđević, 2017; Filipović Đurđević, Đurđević, & Kostić, 2009).

## References


Armstrong, B. C. (2012). *The Temporal dynamics of word comprehension and response selection: Computational and behavioral studies*. (Unpublished doctoral dissertation) of Philosophy). Psychology Department, Carnegie Mellon University, Pennsylvania.

Armstrong, B. C., & Plaut, D. C. (2016). Disparate semantic ambiguity effects from semantic processing dynamics rather than qualitative task differences. *Language, Cognition, and Neuroscience, 1*(7), 1–27. doi:10.1080/23273798.2016.1171366

Armstrong, B. C., Tokowicz, N., & Plaut, D. C. (2012). eDom: Norming software and relative meaning frequencies for 544 English homonyms. *Behavior Research Methods, 44*, 1015–1027. doi:10.3758/s13428-012-0199-8

Azuma, T. (1996). Familiarity and relatedness of word meanings: Ratings for 110 homographs. *Behavior Research Methods, Instruments, and Computers, 28*(1), 109–124. doi:10.3758/BF03203645

Azuma, T., & 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*(4), 484–504. doi:10.1006/jmla.1997.2502

Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research, 32*, 12–28. doi:10.21500/20112084.807

Balota, D. A., Pilotti, M., & Cortese, M. J. (2001). Subjective frequency estimates for 2,938 monosyllabic words. *Memory and Cognition, 29*, 639–647. doi:10.3758/BF03200465

Borowsky, R., & Masson, M. E. J. (1996). Semantic ambiguity effects in word identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(1), 63–85. doi:10.1037//0278-7393.22.1.63

Clark, H. H. (1973). The Language-as-Fixed-Effect-Fallacy: A Critique of Language Statistics in Psychological Research. *Journal of Verbal Learning and Verbal Behavior, 12*, 335–359. doi:10.1016/S0022-5371(73)80014-3

Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance VI* (pp. 535–555). NJ: Erlbaum.

Cover, T. M., & Thomas, J. A. (1991). *Elements of information theory*. New York: John Wiley & Sons. doi:10.1002/0471200611

Eddington, C. M., & Tokowicz, N. (2015). How meaning similarity influences ambiguous word processing: the current state of the literature. *Psychonomic Bulletin & Review, 22*(1), 13–37. doi:10.3758/s13423-014-0665-7

Ferraro, F. R., & Kellas, G. (1990). Normative data for number of word meanings. *Behavior Research Methods, Instruments, and Computers, 22*(6), 491–498. doi:10.3758/BF03204432

Filipović Đurđević, D. (2007). *Polysemy effect in processing of Serbian nouns* (Unpublished doctoral dissertation). University of Belgrade, Serbia.

Filipović Đurđević, D. (2018). Naïve discrimination learning approach to polysemy. *Proceedings of the XXIV Scientific Conference Empirical Studies in Psychology*. March 23-25, 2018, Faculty of Philosophy, University of Belgrade, Institute of Psychology, Laboratory for Experimental Psychology, Faculty of Philosophy, University of Belgrade, 63–64.

Filipović Đurđević, D., Đurđević, Đ., & Kostić, A. (2009). Vector based semantic analysis reveals absence of competition among related senses. *Psihologija, 42*, 95–106. doi:10.2298/PSI0901095F

Filipović Đurđević, D., & Kostić, A. (2008). The effect of polysemy on processing of Serbian nouns. *Psihologija, 41*(1), 69–86. doi:10.2298/PSI0801059F

Filipović Đurđević, D., & Kostić, A. (2017). Number, Relative Frequency, Entropy, Redundancy, Familiarity, and Concreteness of Word Senses: Ratings for 150 Serbian Polysemous Nouns. In S. Halupka-Rešetar & S. Martínez-Ferreiro (Eds.) *Studies in Language and Mind 2* (pp. 13–77). Novi Sad: Filozofski fakultet.

Filipović Đurđević, D., & Kostić, A. We probably sense sense probabilities. *Under review*.

Gawlick-Grendell, L. A., & Woltz, D. J. (1994). Meaning dominance norms for 120 homographs. *Behavior Research Methods, Instruments, and Computers, 26*(1), 5–25. doi:10.3758/BF03204557

Gelman, A., & Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511790942

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. doi:10.1037/0096-3445.113.2.256

Gilhooly, K. J., & Logie, R. H. (1980). Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. *Behavior Research Methods & Instrumentation, 12*(4), 395–427. doi:10.3758/BF03201693

Hino, Y., & Lupker, S. J. (1996). Effects of polysemy in lexical decision and naming: An alternative to lexical access accounts. *Journal of Experimental Psychology: Human Perception and Performance, 22*(6), 1331–1356. doi:10.1037//0096-1523.22.6.1331

Jastrzembski, J. E. (1981). Multiple meanings, number of related meanings, frequency of occurrence, and the lexicon. *Cognitive Psychology, 13*(2), 278–305. doi:10.1016/0010-0285(81)90011-6

Kellas, G., Ferraro, R. R., & Simpson, G. B. (1988). Lexical Ambiguity and the Timecourse of Attentional Allocation In Word Recognition. *Journal of Experimental Psychology: Human Perception and Performance, 14*(4), 601–609. doi:10.1037//0096-1523.14.4.601

Kostić, Đ. (1999). *Frekvencijski rečnik savremenog srpskog jezika* [*Frequency Dictionary of Contemporary Serbian Language*]. Beograd: Institut za eksperimentalnu fonetiku i patologiju govora i Laboratorija za eksperimentalnu psihologiju.

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods, 44*(2), 314–324. doi:10.3758/s13428-011-0168-7

Medeiros, J., & Armstrong, B. C. (2017). *Semantic Ambiguity Effects: A Matter of Time?* In Proceedings of the 39th Annual Conference of the Cognitive Science Society, pp. 2693–2698. Mahwah, N: Lawrence Erlbaum Associates.

Millis, M. L., & Bution, S. B. (1989). The effect of polysemy on lexical decision time: Now you see it, now you don't. *Memory & Cognition, 17*(2), 141–147. doi:10.3758/BF03197064

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. doi:10.1037/h0025327

R Core Team (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/

Rice, C., Beekhuizen, B., Dubrovsky, V., Stevenson, S., & Armstrong, B. C. (2018). A comparison of homonym meaning frequency estimates derived from movie and television subtitles, free association, and explicit ratings. *Behavior Research Methods, 50*, 1399–1425. doi:10.3758/s13428-018-1107-7

RMS – Rečnik Matice Srpske. (1967 – 1976). *Rečnik srpskohrvatskoga književnog jezika, t. I-VI* [*Dictionary of Serbo-Croatian literary language*]. Novi Sad: Matica Srpska.

Rodd, J. M., Gaskell, M. G., & Marslen-Wilson, W. D. (2002). Making sense of semantic ambiguity: Semantic competition in lexical access. *Journal of Memory and Language, 46*, 245–266. doi:10.1006/jmla.2001.2810

Rodd, J. M., Gaskell, M. G., & Marslen-Wilson, W. D. (2004). Modelling the effects of semantic ambiguity in word recognition. *Cognitive Science, 28*, 89–104. doi:10.1207/s15516709cog2801_4

Rubenstein, H., Garfield, L., & Millikan, J. A. (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior, 9*(5), 487–494. doi:10.1016/S0022-5371(70)80091-3

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*(3), 379–423. doi:10.1002/j.1538-7305.1948.tb01338.x

Twilley, L. C., Dixon, P., Taylor, D., & Clark, K. (1994). University of Alberta norms of relative meaning frequency for 566 homographs. *Memory and Cognition, 22*(1), 111–126. doi:10.3758/BF03202766

Van Rij, J., Wieling, M., Baayen, R., & van Rijn, H. (2017). *"itsadug: Interpreting Time Series and Autocorrelated Data Using GAMMs"*. R package version 2.3.

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis.* New York: Springer-Verlag. doi:10.1007/978-3-319-24277-4_9

Wood, S. N. (2006). *Generalized Additive Models: An Introduction with R*. Chapman and Hall/CRC. MY, Florida: Boca Raton. doi:10.1201/9781420010404
Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society B, 731*, 3–36. doi:10.1111/j.1467-9868.2010.00749.x

## Appendix

Table A
*A list of words presented in the norming study and in the experiment accompanied by their descriptions excerpted from the dictionary and frequency dictionary, and collected in the norming study*

| Word | Ambiguity | Type of ambiguity (H – homonymy; P – polysemy) | Lemma frequency (per 2 millions) | Word length in letters | Colthaerts N | Word familiarity | Word concreteness | *H* | *N* | *T* |
|---|---|---|---|---|---|---|---|---|---|---|
| amper | ambiguous | H | 10 | 5 | 1 | 4.87 | 2.05 | 0.14 | 2 | 0.52 |
| bar | ambiguous | H with P | 32 | 3 | 28 | 6.51 | 5.70 | 0.58 | 5 | 0.18 |
| bit | ambiguous | H with P | 26 | 3 | 18 | 5.05 | 2.10 | 0.44 | 4 | 0.27 |
| bob | ambiguous | H | 2 | 3 | 15 | 4.55 | 6.25 | 0.56 | 4 | 0.08 |
| bor | ambiguous | H | 218 | 3 | 18 | 6.53 | 6.85 | 0.33 | 3 | 0.30 |
| dunja | ambiguous | H | 42 | 5 | 5 | 6.49 | 6.65 | 0.35 | 3 | 0.27 |
| džin | ambiguous | H with P | 27 | 4 | 1 | 6.21 | 5.55 | 0.30 | 2 | 0.00 |
| golf | ambiguous | H | 0.1 | 4 | 2 | 6.11 | 5.65 | 0.46 | 4 | 0.24 |
| griz | ambiguous | H with P | 8 | 4 | 7 | 6.13 | 5.55 | 0.30 | 2 | 0.01 |
| herc | ambiguous | H | 0.1 | 4 | 1 | 5.78 | 2.55 | 0.24 | 2 | 0.19 |
| java | ambiguous | H | 379 | 4 | 13 | 5.93 | 2.60 | 0.37 | 3 | 0.23 |
| kubura | ambiguous | H with P | 6 | 6 | 2 | 3.84 | 6.45 | 0.32 | 3 | 0.32 |
| kup | ambiguous | H with P | 38 | 3 | 15 | 5.48 | 4.70 | 0.29 | 2 | 0.05 |
| lala | ambiguous | H but also a verb | 8 | 4 | 21 | 6.31 | 6.55 | 0.44 | 4 | 0.27 |
| lira | ambiguous | H with P | 54 | 4 | 10 | 4.32 | 6.25 | 0.28 | 2 | 0.07 |

| Word | Ambiguity | Type of ambiguity (H – homonymy; P – polysemy) | Lemma frequency (per 2 millions) | Word length in letters | Colthaerts N | Word familiarity | Word concreteness | *H* | *N* | *T* |
|---|---|---|---|---|---|---|---|---|---|---|
| mah | ambiguous | H with P | 92 | 3 | 15 | 4.19 | 2.60 | 0.16 | 2 | 0.47 |
| mig | ambiguous | H with P | 11 | 3 | 9 | 5.00 | 4.55 | 0.28 | 2 | 0.07 |
| mina | ambiguous | H with P | 41 | 4 | 16 | 6.11 | 6.25 | 0.47 | 3 | 0.01 |
| mišić | ambiguous | H | 53 | 5 | 1 | 6.59 | 6.05 | 0.35 | 3 | 0.28 |
| mladež | ambiguous | H | 11 | 6 | 0 | 6.54 | 6.20 | 0.29 | 2 | 0.04 |
| mol | ambiguous | H | 22 | 3 | 11 | 5.01 | 2.65 | 0.52 | 4 | 0.14 |
| nana | ambiguous | H with P | 26 | 4 | 12 | 6.57 | 6.30 | 0.40 | 3 | 0.16 |
| osa | ambiguous | H with P | 13 | 3 | 7 | 6.53 | 6.00 | 0.30 | 2 | 0.00 |
| otok | ambiguous | H with P | 91 | 4 | 2 | 6.01 | 5.95 | 0.30 | 2 | 0.00 |
| paša | ambiguous | H with P | 53 | 4 | 13 | 5.00 | 4.50 | 0.39 | 3 | 0.19 |
| peć | ambiguous | H only by personal name | 177 | 3 | 6 | 6.32 | 6.55 | 0.24 | 2 | 0.19 |
| pijuk | ambiguous | H with P | 29 | 5 | 2 | 3.60 | 5.20 | 0.29 | 2 | 0.02 |
| pik | ambiguous | H with P | 8 | 3 | 15 | 4.36 | 3.60 | 0.40 | 3 | 0.16 |
| pop | ambiguous | H with P | 40 | 3 | 7 | 6.65 | 6.20 | 0.37 | 3 | 0.23 |
| posuda | ambiguous | H with P | 12 | 6 | 6 | 6.84 | 6.15 | 0.23 | 2 | 0.24 |
| prag | ambiguous | H with P | 386 | 4 | 7 | 6.66 | 6.60 | 0.29 | 2 | 0.02 |
| rak | ambiguous | P | 28 | 3 | 23 | 6.68 | 6.10 | 0.46 | 3 | 0.04 |
| ram | ambiguous | H | 24 | 3 | 17 | 6.23 | 6.50 | 0.30 | 2 | 0.02 |
| remi | ambiguous | H | 35 | 4 | 2 | 5.43 | 3.45 | 0.39 | 3 | 0.18 |
| reper | ambiguous | H | 0.1 | 5 | 1 | 6.36 | 5.60 | 0.30 | 2 | 0.00 |
| ris | ambiguous | H | 10 | 3 | 11 | 5.70 | 6.50 | 0.32 | 3 | 0.33 |
| rizla | ambiguous | H | 0.1 | 5 | 0 | 4.75 | 6.40 | 0.18 | 2 | 0.40 |
| rok | ambiguous | H with P | 241 | 3 | 19 | 6.84 | 2.80 | 0.39 | 3 | 0.19 |

| Word | Ambiguity | Type of ambiguity (H – homonymy; P – polysemy) | Lemma frequency (per 2 millions) | Word length in letters | Colthaerts N | Word familiarity | Word concreteness | *H* | *N* | *T* |
|---|---|---|---|---|---|---|---|---|---|---|
| šljaka | ambiguous | H but also a verb | 4 | 6 | 1 | 5.48 | 5.45 | 0.38 | 3 | 0.20 |
| sud | ambiguous | H with P | 221 | 3 | 14 | 6.48 | 4.15 | 0.29 | 2 | 0.03 |
| šut | ambiguous | H with P | 22 | 3 | 13 | 5.73 | 4.90 | 0.26 | 2 | 0.14 |
| tuš | ambiguous | H with P | 3 | 3 | 11 | 6.78 | 6.20 | 0.30 | 2 | 0.01 |
| vlada | ambiguous | H but also a verb | 1452 | 5 | 5 | 6.67 | 3.20 | 0.44 | 3 | 0.08 |
| žal | ambiguous | H | 148 | 3 | 11 | 4.77 | 3.05 | 0.15 | 2 | 0.51 |
| žiža | ambiguous | H only by personal name | 23 | 4 | 5 | 4.38 | 3.50 | 0.21 | 3 | 0.56 |
| zrak | ambiguous | H with P | 512 | 4 | 10 | 5.93 | 4.95 | 0.29 | 2 | 0.04 |
| ćerka | unambiguous | unambiguous | 55 | 5 | 5 | 6.85 | 5.85 | | | |
| afera | unambiguous | unambiguous | 9 | 5 | 1 | 6.25 | 2.25 | | | |
| anatom | unambiguous | unambiguous | 1 | 2 | 0 | 3.55 | 4.60 | | | |
| april | unambiguous | unambiguous | 453 | 2 | 0 | 6.85 | 2.45 | | | |
| autor | unambiguous | unambiguous | 62 | 5 | 1 | 6.50 | 4.55 | | | |
| azbest | unambiguous | unambiguous | 2 | 6 | 0 | 3.30 | 4.95 | | | |
| bakar | unambiguous | unambiguous | 43 | 5 | 4 | 5.70 | 6.00 | | | |
| barut | unambiguous | unambiguous | 30 | 5 | 0 | 5.45 | 6.15 | | | |
| berber | unambiguous | unambiguous | 4 | 6 | 4 | 3.90 | 6.05 | | | |
| brak | unambiguous | unambiguous | 47 | 4 | 13 | 6.60 | 3.10 | | | |
| bu? | unambiguous | unambiguous | 2 | 3 | 8 | 6.40 | 6.30 | | | |
| bulka | unambiguous | unambiguous | 29 | 5 | 3 | 4.45 | 6.30 | | | |
| duet | unambiguous | unambiguous | 27 | 4 | 1 | 6.10 | 3.70 | | | |
| fes | unambiguous | unambiguous | 2 | 3 | 7 | 3.35 | 6.15 | | | |

| Word | Ambiguity | Type of ambiguity (H – homonymy; P – polysemy) | Lemma frequency (per 2 millions) | Word length in letters | Colthaerts N | Word familiarity | Word concreteness | *H* | *N* | *T* |
|---|---|---|---|---|---|---|---|---|---|---|
| grb | unambiguous | unambiguous | 18 | 3 | 5 | 6.15 | 5.75 | | | |
| hektar | unambiguous | unambiguous | 72 | 2 | 2 | 4.70 | 2.60 | | | |
| hektar | unambiguous | unambiguous | 72 | 6 | 2 | 5.15 | 2.60 | | | |
| hir | unambiguous | unambiguous | 14 | 3 | 15 | 6.60 | 1.80 | | | |
| hotel | unambiguous | unambiguous | 120 | 5 | 0 | 6.40 | 6.60 | | | |
| jul | unambiguous | unambiguous | 279 | 3 | 6 | 6.75 | 2.45 | | | |
| jun | unambiguous | unambiguous | 342 | 3 | 11 | 6.80 | 2.55 | | | |
| koks | unambiguous | unambiguous | 8 | 2 | 5 | 4.15 | 6.00 | | | |
| korov | unambiguous | unambiguous | 58 | 5 | 3 | 6.30 | 6.20 | | | |
| kujna | unambiguous | unambiguous | 109 | 5 | 1 | 6.05 | 6.20 | | | |
| laž | unambiguous | unambiguous | 192 | 3 | 9 | 6.80 | 2.00 | | | |
| lift | unambiguous | unambiguous | 11 | 2 | 3 | 6.50 | 7.00 | | | |
| lim | unambiguous | unambiguous | 23 | 2 | 15 | 5.90 | 6.50 | | | |
| mač | unambiguous | unambiguous | 175 | 3 | 15 | 6.20 | 6.90 | | | |
| mart | unambiguous | unambiguous | 472 | 4 | 4 | 6.85 | 2.50 | | | |
| nafta | unambiguous | unambiguous | 46 | 5 | 2 | 6.20 | 5.70 | | | |
| nar | unambiguous | unambiguous | 21 | 3 | 21 | 5.40 | 6.60 | | | |
| opal | unambiguous | unambiguous | 7 | 4 | 6 | 3.55 | 5.45 | | | |
| papaja | unambiguous | unambiguous | 0.1 | 2 | 0 | 4.55 | 6.60 | | | |
| plik | unambiguous | unambiguous | 12 | 4 | 5 | 5.80 | 6.60 | | | |
| ponoć | unambiguous | unambiguous | 270 | 5 | 3 | 6.60 | 3.50 | | | |
| srp | unambiguous | unambiguous | 61 | 3 | 8 | 4.65 | 6.30 | | | |
| tepih | unambiguous | unambiguous | 48 | 5 | 0 | 6.45 | 6.90 | | | |
| tundra | unambiguous | unambiguous | 2 | 6 | 0 | 4.00 | 5.70 | | | |
| varjača | unambiguous | unambiguous | 1 | 2 | 0 | 6.00 | 6.95 | | | |

| Word | Ambiguity | Type of ambiguity (H – homonymy; P – polysemy) | Lemma frequency (per 2 millions) | Word length in letters | Colthaerts N | Word familiarity | Word concreteness | *H* | *N* | *T* |
|---|---|---|---|---|---|---|---|---|---|---|
| vino | unambiguous | unambiguous | 574 | 4 | 4 | 6.85 | 6.80 | | | |
| violina | unambiguous | unambiguous | 103 | 2 | 0 | 6.55 | 6.90 | | | |
| vrabac | unambiguous | unambiguous | 46 | 6 | 1 | 6.75 | 6.85 | | | |
| zajam | unambiguous | unambiguous | 143 | 5 | 4 | 6.40 | 3.60 | | | |
| žbun | unambiguous | unambiguous | 38 | 4 | 1 | 6.05 | 6.75 | | | |
| zet | unambiguous | unambiguous | 25 | 3 | 9 | 6.15 | 5.75 | | | |
| žir | unambiguous | unambiguous | 11 | 3 | 16 | 5.85 | 6.90 | | | |

*Note. H* - entropy of the meaning frequency distribution; *N* - number of meanings; *T* -redundancy of the meaning frequency distribution.

**Dušica Filipović Đurđević**

Odeljenje za psihologiju, Filozofski fakultet, Univerzitet u Beogradu

Laboratorija za eksperimentalnu psihologiju, Filozofski fakultet, Univerzitet u Beogradu

Laboratorija za eksperimentalnu psihologiju, Filozofski fakultet, Univerzitet u Novom Sadu

# UJEDNAČENOST VEROVATNOĆA ZNAČENJA I OBRADA HOMONIMIJE U SRPSKOM JEZIKU

U ovom istraživanju grupa homonimnih imenica srpskog jezika (imenica sa višestrukim nepovezanim značenjima) izlagana je u normativnoj studiji i u eksperimentu sa zadatkom vizuelne leksičke odluke. Govornici, kojima je srpski jezik maternji, navodili su značenja homonima i procenjivali reči na skali familijarnosti i konkretnosti. Na osnovu njih, formirana je prva baza homonima srpskog jezika, koja sadrži značenja homonimnih imenica srpskog jezika koja su poznata ispitanicima, kao i procenjene verovatnoće svakog značenja, broj značenja, redundansu i entorpiju distribucije verovatnoće značenja, poznatost reči i konkretnost. U zadatku vizuelne leksičke odluke ponovljen je nalaz o sporijoj obradi homonima u odnosu na jednoznačne reči. Dodatno, obrada homonimnih imenica dovedena je u vezi sa redundansom – informaciono-teorijskom merom koja opisuje ujednačenost verovatnoća značenja. Rezultati su pokazali da su homonimi sa većom redundansom distribucije verovatnoća značenja (tj. neujednačenim verovatnoćama značenja) imali kraće vreme prepoznavanja. Ovaj nalaz je u skladu sa predikcijom izvedenom iz pristupa obradi višeznačnih reči koji se zasniva na dinamici razrešavanja značenja, po kojoj kompeticija između nepovezanih značenja dovodi do sporije obrade homonima. Međutim, u slučaju broja značenja i entropije, obrazac rezultata je donekle odstupao, zbog čega je potrebno nastaviti sa itraživanjem obrade višeznačnih reči.

**Ključne reči**: broj značenja, entropija, homonimija, redundansa, zadatak vizuelne leksičke odluke