

# Factors of Collocation Attainment among Native Croatian Language Speakers

Antonia Ordulj<sup>1</sup> and Lidija Cvikić<sup>2</sup>

<sup>1</sup>*Croaticum – Center for Croatian as a Second and Foreign Language, Faculty of Humanities and Social Sciences, University of Zagreb*

<sup>2</sup>*Faculty of Teacher Education, University of Zagreb*

## Abstract

*Research on collocations in the Croatian language has mostly been directed towards issues relating to linguistics and lexicography, while research on their use is scarce and mostly directed towards second language speakers (Burić & Lasić, 2012; Ordulj, 2016; Petrović, 2007). This paper brings forth the results of research on the use of collocations among native speakers of the Croatian language. Based on the corpus analysis of characteristics of collocations and investigation of native Croatian language speakers, the influence of two factors on the use of collocations was researched: frequency of collocations and associative strength of collocation elements. The use of collocations in a morphologically unmarked form (nominative) and in dependent cases, along with their use in the narrower and broader context was researched. The research results indicate that the key factors for using collocations by native speakers are factors of frequency and associative strength, while the morphological form of the collocation is irrelevant to native speakers. The research also showed stronger influence of context on the use of less frequent collocations and weaker associative strength from the use of collocations of high frequency and strong associative strength. The results obtained are relevant for understanding general processes in the acquisition of the Croatian language as the mother tongue and suggest a starting point for considering more efficient ways of teaching the Croatian language.*

**Key words:** *associative strength of collocation; context; frequency; mother tongue attainment.*

## Introduction

Numerous linguistic research from the onset of the 1950s emphasize collocations as a neglected constituent of lexical competence of a second language (Gitsaki, 1999; Petrović, 2007). However, over the last twenty years, the mentioned lexical phenomenon has been receiving more and more attention (Bahns & Eldaw, 1993; Gitsaki, 1999; Jaén, 2007; Koya, 2003; Martyńska, 2004). Papers dealing with collocational competence in the mother tongue are sporadic firstly due to the ingrained opinion that native speakers intuitively combine lexical items as they are constantly exposed to and use the mother tongue (cf. Borić, 2002). Although collocations in Croatian linguistics seem to be attracting more attention, primarily for English as a second language, research relating to the Croatian language continues to be sporadic. To date, the most detailed description of noun collocations in Croatian as the mother tongue (L1) was researched by Blagus Bartolec (2014) while collocations in Croatian as a second language (CSL) was the subject of study by Petrović (2007), and Burić and Lasić (2012), while Ordulj (2016) is the first to present a detailed analysis of collocations as a constituent of lexical competence in Croatian as a second language.

The term collocation is most frequently defined as co-occurrence of particular lexical items at the syntagmatic level (Borić, 2004; Firth, 1957; Lewis, 2000; Nation, 2000; Petrović, 2007). In that way, the meanings of particular lexical items are combined for achieving comprehensive meaning. The concept itself was introduced into linguistics by John R. Firth (1957), who speaks of collocation solely from the perspective of lexical relationships and lexis as an independent level of study, detached from grammar. Collocations should be differentiated from other similar word associations: phrasemes, names and terms. Turk (2000) cautions about their similarities and differences. Among similarities, Turk (2000) includes complex structure, semantic conversion of at least one item and productivity, while stability, connotative power, idiomaticity and semantic relationships are marked as differences. An important characteristic in differentiating the mentioned lexical combinations is their communicative function. Phrasemes are expressions of secondary nominations and point to personal, emotional and judgmental attitudes, where names serve for communication and definition of new concepts among professionals from scientific or professional area while the primary task of terms is to identify an object. Blagus Bartolec (2014, p. 80) states that collocation is a “particular lexical combination at the syntagmatic level based on the semantic connectedness of individual lexical items which establishes their meanings.” Blagus Bartolec (2014) restricts a collocation to two content words which formally denote a syntagm.

Research on collocations can be classified with respect to the various approaches in studying collocations: lexical-composition (Firth, 1957; Halliday, 1966; Sinclair, 1966), semantic-syntactic (Cruse, 1986; Katz & Fodor, 1963; Mel'čuk, 2008) and structural approach (Greenbaum, 1974; Matthews, 1964; Mitchell, 1971). Representatives of the lexical-composition approach consider lexis an autonomous entity independent

of grammar. Firth (1957) is the first to introduce the concepts collocability and collocational meaning and attempts to explain the term collocation. He defines collocation as a manner of realizing lexical meaning which stems from the usual co-occurrence of two lexemes along the syntagmatic axis (Firth, 1957). Firth's theoretical considerations are further elaborated by his students Halliday (1966) and Sinclair (1966). Although they did not offer answers to fundamental questions regarding collocations, Firth (1957), Halliday (1966) and Sinclair (1966) are instrumental in the introduction of the term collocation, work on terminology and the analysis of semantic relationships between lexical items at the syntagmatic level. The semantic-syntactic approach (Cruse, 1986; Katz & Fodor, 1963; Melčuk, 2008) attempts to predict the collocational span of particular lexical item based on semantic characteristics. This approach was not successful in explaining semantic relationships between collocational combinations, so greater emphasis is placed on studying paradigmatic lexical relationships, leaving collocations aside and treating them as exceptions to the rule. The structural approach advocates the unity of lexis and grammar in trying to establish a collocational structure. Research by representatives of the structural approach (Greenbaum, 1974; Matthews, 1964; Mitchell, 1971) are seen as attempts to explain collocations using structural patterns taking into consideration collocational limitations. Despite numerous research and theoretical approaches, there are still no answers to basic questions regarding collocations, i.e. why certain lexis collocate and based on what criteria, how to clearly differentiate collocations from other lexical combinations and how meaning of collocation is achieved.

In Croatian reference literature, collocations are approached from various aspects. Thus, we can speak of papers dealing with the contrastive, semantic and lexicographic analysis of collocations and collocations in language for specific purposes. What is more, there is a noticeable interest in the glottodidactic aspect of researching collocations, which emphasizes the development of the collocational competence where learning and teaching lexical combinations which co-occur at the syntagmatic level is placed in the forefront.

Papers dealing with the contrastive and semantic analysis of collocations (Borić, 1998; 1996; Gulešić Machata & Machata, 2007) firstly speak of the semantic relationships between collocational elements and the emergence of collocational meaning, along with the issue of translation equivalence of collocation. Within semantic discussions, we single out the paper by Blagus Bartolec (2014) considering its attempt to define and describe noun collocations in the Croatian language for the first time and the emergence of collocational meaning through syntagmatization. Collocations in the Croatian language originate through syntagmatization, which means that they do not have the status of lexical units as semantic adaptation occurs under the influence of the semantic potential of the second word entering the collocational relationship. Owing to the semantic potential of the collocational components, there is a modification of one element by the other and formation of a new, collocational meaning which is

“motivated by an extralinguistic phenomenon that is being referred to (e.g. *sunny/rainy day*, *long/short hair*) (Blagus Bartolec, 2014 p. 85)”. Furthermore, Blagus Bartolec (2014) specifies, for the first time, the division of collocations according to type with respect to formal and semantic features of lexical units which co-occur. Numerically, most representative papers are those dealing with the lexicographical analysis of collocations in Croatian dictionaries (Borić, 2002; Duplančić Rogošić, 2007; Mihaljević, 1991; Petrović, 2008a, 2008b; Pritchard, 1998; Turk, 2010) considering that the issue of whether collocations should be included in dictionaries and the fact that there is no Croatian dictionary of collocations is still unresolved. Interest for collocations in language for specific purposes, primarily English, is shown by a growing number of authors attempting to point to the issue considering the distinctiveness of the lexis of each profession. The most frequent papers on collocations in language for specific purposes are in the medical (Miščin, 2012; Miščin & Pavičić Takač, 2013; Štefić, Mravak–Stipetić, & Borić, 2010) and technical profession (Kos, 2012; Špiranec, 2005).

As mentioned, with the exception of several papers on Croatian as a second language (Burić & Lasić, 2012; Ordulj, 2016) Croatian linguistics in general is lacking in research on the development of collocational competence. Firstly, it is important to research the development of the collocational competence in Croatian as a native language as the use of collocations can be a significant indicator of the degree of lexical knowledge. Additionally, with the description of collocational competence in Croatian as the first language, a target level of attainment of Croatian as a second language can be defined. Therefore, the main aim of this paper is to research factors which affect the collocational competence of native speakers of the Croatian language so as to obtain a clearer picture of attainment of collocations in Croatian.

### ***Collocational Competence***

Collocational competence is one of the most important segments of productive and receptive lexical knowledge considering that collocations enable native and non-native speakers to differentiate between the meaning variants of polysemic words and to understand connotative meaning (Henriksen, 2013). Furthermore, collocations contribute to the fluency and natural communication and narrow the meaning of particular lexical units that enter the collocation relationship creating associative relationships between collocational elements (Borić, 2004; Henriksen, 2013; Nation, 2001). Many applied linguists have been indicating the significance of collocations, particularly when referring to second language acquisition, emphasizing how collocations in foreign language teaching receive little attention and should be approached more systematically and taught more explicitly (Bahns & Eldaw, 1993). It is generally assumed that with the development of general vocabulary collocational competence is also developed (Bonci, 2002; Gitsaki, 1999; Koya, 2003). At the same time, students at higher language levels show better knowledge of lexical

and grammatical collocations in the second language (Begagić, 2014; Bonci, 2002; Gitsaki, 1999). Several factors influence the development of collocational competence. Among the most significant ones are frequency of exposure to collocations, differences in the mother tongue, structural and semantic complexity of collocation types, and educational environment, i.e. exposure to various collocation types through the teaching process (Begagić, 2014; Gitsaki, 1999).

In addition to papers dealing exclusively with collocational competence in the second language, recent papers deal with the comparison of the estimation of frequency of collocations among native and non-native speakers. Research by Siyanova-Chanturie and Spine (2015) shows that native and non-native speakers of the Italian language do not differ in their estimation of frequent collocations and their estimate is positively correlated with the frequency of collocations from the corpus. In the same research, native speakers were more successful in the estimate of collocations of low frequency, as was the advanced group of speakers of Italian as a second language. However, the greatest problem for both groups of speakers were collocations of medium frequency as they occur in the middle of the continuum. In second language research, emphasis is given to highly frequent language units because they are frequent and therefore useful for communication, along with low-frequency language units, as it is highly unlikely that language students will be exposed to them in the language input, so they have to be explicitly taught. Siyanova and Schmitt (2008) attempted to establish the extent to which non-native speakers use appropriate collocations in written production and whether they are successful in the estimate of the degree of frequency of collocations compared to native English speakers. The research confirmed that non-native speakers use frequent and associatively strong collocation in the same amount as native speakers.

Therefore, the research that is being put forward in this paper is directed to those two key factors in the use of collocations: frequency and associative strength of collocational elements.

## **Methods**

The basic aim of the paper is to examine applied knowledge of noun collocations with respect to frequency and associative strength of collocational elements in Croatian as the native language.

The following research objectives are set forth:

1. Examine the knowledge of noun collocations with respect to the frequency and strength of collocational elements.
2. Examine applied knowledge of morphologically unmarked collocations in the nominative and morphologically marked collocations in the oblique cases with respect to associative strength of collocational elements and frequency of collocations.
3. Examine the productive knowledge of noun collocations in a broader context.

Considering the research objectives, the following hypotheses were set:

H1: Participants will have better knowledge of collocations of greater frequency and stronger associative strength than those of low frequency and weaker associative strength.

H2: Participants will show equal accuracy in answers for morphologically unmarked collocations in the nominative and morphologically marked collocations in oblique cases, regardless of the frequency and associative strength of the collocations.

H3: For collocations in the broader context, participants will show equal accuracy of answers for collocations of higher frequency and stronger associative strength.

Research on collocational competence was carried out in two phases. The first phase demanded a corpus research of collocations in the hrWaC corpus (Croatian Web Corpus, Ljubešić & Erjavec, 2011; hrWaC) followed by an examination of the associative strength of collocation items. The second phase was marked by the development of an instrument, and research with participants, native speakers of Croatian language was carried out.

### ***Selection of Collocations for Research Purposes***

In order to carry out the research, several hundred collocations had to be collected representing the working corpus and were used in developing the instruments. The research is dominated by collocations of the basic type where the primary meaning is realized, i.e. as Blagus Bartolec (2014, p. 93) states there is a “basic level of primary meaning specification.” In the corpus analysis, on which the selection of collocations was based, the binary division of collocation elements on the base and the collocator was accepted, where the base was a noun and the collocator an adjective. The collocations were selected based on frequency (higher and lower) and associative strength of collocational elements (stronger and weaker collocations).

For research with native speakers the same testing material was selected as for research carried out by Ordulj (2016) on non-native speakers of Croatian language for which collocations were drawn from essays written by non-native speakers of the Croatian language (for more detail see Ordulj, 2016). In total, 300 essays were analyzed and they make up the first Student corpus of Croatian as a second language – CROLTEC (Mikelić Preradović, Berać, & Boras, 2015), which is still not publicly available.

The validity of collocations was checked in the Large Dictionary of the Croatian Language (Anić, 2003) and the hrWaC corpus (Ljubešić & Erjavec, 2011), which currently represents the greatest corpus for Croatian language comprising 1.9 billion tokens. The hrWaC corpus is the first web corpus for the Croatian language comprising all web pages from the national domain of the Republic of Croatia (.hr). In this research, the search for collocations in the hrWaC corpus was done using the CQL programming language (*Corpus Query Language*), which enables searches of

the corpus with marked words with respect to default attribute. For the purpose of this paper, collocations were searched according to the attribute *word*, and the CQL query for the collocation *primary school* looks like this: [word="prim.\*" & tag="A.\*"] [word="schoo.\*" & tag="N.\*"]

Using computer language tools in the hrWaC corpus, the frequency of collected collocations was calculated and, according to frequency, they were divided into two groups - higher (HF) and lower (LF) frequency. Such a division is in accordance with Sinclair (1991), who states that the frequency of searched terms depends of the type and number of tokens within a corpus, and the larger the corpus the more reliable the results are.

### **Participants**

The research on collocational competence in Croatian as the mother tongue was carried out during the winter semester of the 2015/2016 academic year. Fifty students of the Faculty of Humanities and Social Sciences in Rijeka participated in research on examining the associative strength of collocations, while seventy students of the Faculty of Teacher Education in Zagreb participated in the research on collocational competence (F=65, M=5). Participants in both research studies are native Croatian language speakers.

### **Instruments**

The research applied two instruments developed for the research purposes. One instrument examined the evaluation of associative strength of collocations, while the other examined the collocational competence of native speakers.

The associative strength of collocational elements was evaluated on a scale from 1 to 5 and the collocations were randomly arranged. Descriptive data were calculated for each collocation (M, SD) representing the average associative strength. Although the evaluations are subjective, the averaged data serve as objective indicators of associative strength, i.e. the strength between the collocational elements. Considering the value of the arithmetic mean (M) after the research, collocations were grouped into two categories – strong (SAS) and weak associative strength (WAS).

Collocational competence was examined using the instrument comprising three fill-in-the-blank tasks: a) task with collocations in morphologically unmarked nominative case; b) task with collocations in oblique cases; c) task with collocations in the broader context. The first task had twenty sentences where the collocations were in the nominative (task A), while in the second task the same collocations were in oblique cases (task D) (Table 1). All tasks contained collocations with respect to the characteristics researched: frequency and high associative strength. In other words, the tasks contained collocations of high frequency and strong associative strength (HF + SAS), high frequency and weak associative strength (HF + WAS), low frequency and strong associative strength (LF + SAS) and low frequency and weak associative strength (LF + WAS).



Table 1

*Distribution of collocations according to frequency and associative strength in tasks with collocations in the nominative (A) and oblique cases (D)*

| FILL IN THE BLANK TASK (A/D)  |                          |
|-------------------------------|--------------------------|
| HF + SAS                      | LF + SAS                 |
| 1. HUMAN RIGHTS               | 1. SYNCHRONIZED SWIMMING |
| 2. SOCIAL NETWORK             | 2. SINGLE PARENT         |
| 3. CULTURAL HERRITAGE         | 3. ORGANIC FOOD          |
| 4. CREDIT CARD                | 4. MASS PRODUCTION       |
| 5. FARMACEUTIC INDUSTRY       | 5. ACTION MOVIE          |
| HF + WAS                      | LF + WAS                 |
| 1. OFFICIAL LANGUAGE          | 1. STATUS SYMBOL         |
| 2. HEALTH CARE                | 2. FIGURE SKATING        |
| 3. SPORTS PAPER               | 3. HAND BAG              |
| 4. NONGOVERNMENT ORGANIZATION | 4. VIRTUAL FRIEND        |
| 5. PERSONAL DATA              | 5. STEREOTYPE ROLE       |

The third task had the collocations inserted in an informative, broader context (task C). It contained ten collocations in oblique cases from task D and ten new collocations which are not present in the previous tasks (Table 2).

Table 2

*Distribution of collocations according to frequency and associative strength in the task with collocations in the broader context (C)*

| FILL IN THE BLANK TASK IN THE BROADER CONTEXT (C) |                      |
|---|----------------------|
| HF + SAS  | LF + SAS             |
| 1. HUMAN RIGHTS                                   | 1. SINGLE PARENT     |
| 2. SOCIAL NETWORK                                 | 2. ORGANIC FOOD      |
| 3. FARMACEUTIC INDUSTRY                           | 3. MASS PRODUCTION   |
| 4. EATING HABITS                                  | 4. PELAGIC FISH      |
| 5. LIFE STANDARD                                  | 5. DEADLY DISEASE    |
| HF + WAS  | LF + WAS             |
| 1. HEALTH CARE                                    | 1. STATUS SYMBOL     |
| 2. NONGOVERNMENT ORGANIZATION                     | 2. VIRTUAL FRIEND    |
| 3. RURAL AREA                                     | 3. ALL-PURPOSE FLOUR |
| 4. CLIMATE CONDITIONS                             | 4. TOURIST CENTER    |
| 5. PURCHASING POWER                               | 5. HOME BUDGET       |

The research covered a uniform number of collocations with respect to the criteria of frequency and strength of associative strength of collocational elements. All tasks had the collocation base (noun), and the subjects had to show productive knowledge of the collocator. Those answers anticipated by the authors of the tests, i.e. this research were considered correct. All other answers were grouped into different types of answers in the selection of collocates of the target collocation with respect to semantic, formational, grammatical or orthographical similarity or difference.

The results were analyzed quantitatively and qualitatively (typology of the type of answer in the selection of collocates). The statistical package Statistica was used for data analysis.



## Research Procedure

The participants in the research were familiar with the research purpose and their participation was voluntary and anonymous. Considering the specificity of the testing material, i.e. the fact that task A (nominative) and task D (oblique cases) contained the same collocations and that task C (collocations in the broader context) shared ten equal collocations with task D, the research within each group was organized according to a predefined scheme, and the order of collocations in all tasks was varied and each participant had a password. The testing materials were distributed individually.

In order to examine the influence of morphological characteristics of words on knowledge of collocations, the participants were divided within groups into two even sub-groups. The first group solved task A with collocations in the nominative, the second solved task D with collocations in oblique cases. The division of tasks within each group was carried out according to the schedule shown in Table 3.

Table 3  
*Distribution of participants within the group according to the type of task*

|         | participant | task | task | task |
|---------|-------------|------|------|------|
| GROUP 1 | 1           | A    | B    | C    |
|         | 2           | C    | B    | A    |
|         | 3           | A    | B    | C    |
|         | 4           | C    | B    | A    |
|         | 5           | A    | B    | C    |
|         | 6           | C    | B    | A    |
| GROUP 2 | 7           | C    | B    | D    |
|         | 8           | D    | B    | C    |
|         | 9           | C    | B    | D    |
|         | 10          | D    | B    | C    |
|         | 11          | C    | B    | D    |
|         | 12          | D    | B    | C    |

The subsequent paragraphs bring forth the results of the research analysis of collocational competence in Croatian as mother tongue (L1)<sup>1</sup>. General results are presented first, followed by more specific results with respect to collocation characteristics (HF, LF, SAS, WAS) and tasks. Furthermore, for better transparency and understanding, the interpretation of results is given immediately following the data analysis.

<sup>1</sup> The research conducted used one task to examine receptive knowledge of collocations with the aim to test the appropriateness of the instrument for examining non-native speakers. Considering that the accuracy of answers was exceptionally high (for collocations of higher frequency  $M = .99$ ;  $SD = .03$ , and for collocations of lower frequency  $M = .97$ ;  $SD = .06$ ), these answers were not analyzed in this paper. Still, it should be emphasized that a significant difference in the accuracy of answers was revealed -  $t(69) = 2.189$ ,  $p < .05$  considering collocation frequency.

## Results and Discussion

The analysis of correct answers for tasks at the productive level will be shown first for tasks with collocations in the nominative and then with collocations in oblique cases, followed by productive knowledge of collocations in the broader context.

### ***Analysis of Correct Answers in Tasks with Collocations in the Nominative and Oblique Cases***

With the aim to examine the accuracy of answers in tasks of productive knowledge of morphologically unmarked collocations in the nominative (Task A) and morphologically marked collocations in oblique cases (Task D) a three-way analysis of variance was conducted with repeated measures on frequency factors (higher, lower) and associative strength (weak, strong) of collocations and independent groups on the factor word form (nominative, oblique case). The results are shown in Table 4. The parameters reported in the table are as follows: degrees of freedom (df), F-value (F), Mean Square Error (MSE) and p-value (p).

Table 4

*Results of the three-way analysis of variance with repeated measures on factors of frequency (higher, lower) and associative strength (weak, strong) of collocations and independent groups on the factor word form (nominative, oblique case) for the variable accuracy of answer*

| Source of variability                        | df   | F      | MSE  | p    |
|--|------|--------|------|------|
| Word form                                    | 1,68 | .018   | .072 | .894 |
| Frequency                                    | 1,68 | 87.356 | .032 | .000 |
| Associative strength                         | 1,68 | 64.889 | .031 | .000 |
| Word form * frequency                        | 1,68 | .041   | .032 | .841 |
| Word form * associative strength             | 1,68 | .225   | .031 | .637 |
| Frequency * associative strength             | 1,68 | 28.015 | .023 | .001 |
| Word form * frequency * associative strength | 1,68 | 1.404  | .023 | .240 |

The main effects of frequency and associative strength of collocations were obtained and a significant interaction of frequency and associative strength of collocations on the accuracy of answers in tasks of productive knowledge among native Croatian speakers was established. Duncan's post hoc test established that the accuracy of answers was highest among collocations of higher frequency and strong associative strength ( $M = .75$ ;  $SE = .02$ ) and it significantly differs in comparison with all remaining collocations (all  $p < .01$ ). Furthermore, the equal level of accuracy was obtained among collocations of higher frequency and low associative strength ( $M = .67$ ;  $SE = .02$ ) and lower frequency and strong associative strength ( $M = .65$ ;  $SE = .02$ ). Accuracy of answers is lowest with collocations of low frequency and weak associative strength ( $M = .38$ ;  $SE = .03$ ). The difference is significant in comparison with all other collocations (all  $p < .001$ ) (Figure 1). The results confirm hypotheses H1 and H2.

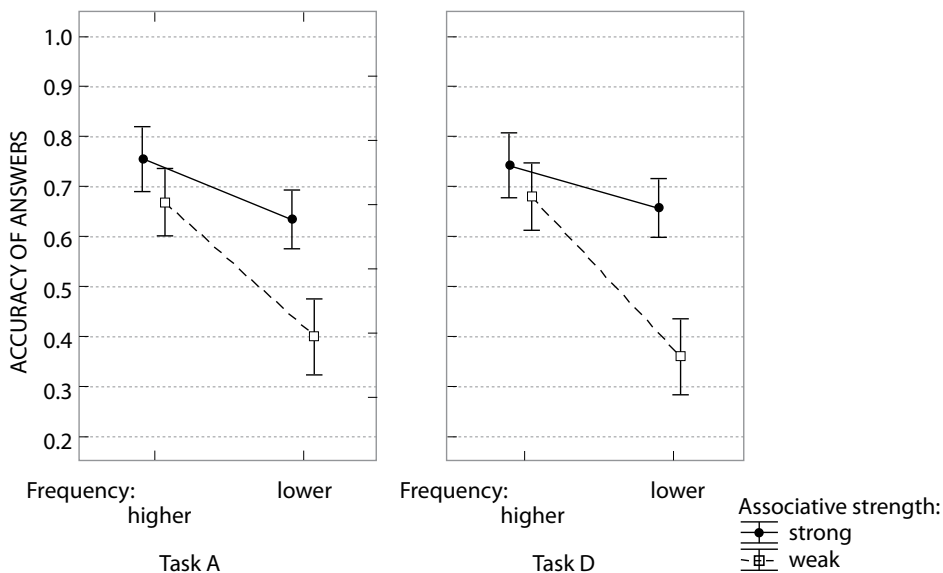


Figure 1. Accuracy of answers for tasks of productive knowledge of morphologically unmarked (Task A) and marked collocations (Task D) with respect to frequency (low, high) and associative strength (weak, strong) of collocational elements

The results obtained show that native speakers gave accurate answers for collocations of high frequency and strong associative strength. That result can be compared to the results of the research by Siyanova and Schmitt (2008) which showed that non-native speakers come closest to native speakers in the use of collocations of high frequency and strong associative strength. Such results are in accordance with knowledge of applied theories of language acquisition (see Cvikić, 2016) according to which different types of frequency (e.g. frequency of token, frequency of paradigm, etc.) are important factors in attaining particular language units or patterns. Therefore, further research should attempt to examine how the acquisition of collocations is influenced by the frequency of the collocation itself with respect to the frequency of its elements. The results indicated a significant difference in the development of collocational competence of native and non-native speakers of Croatian. Research conducted by Ordulj (2016) showed that for non-native speakers at lower levels of attainment of Croatian, the production of collocations is influenced solely by the factor of frequency and not associative strength, while in the production of collocations by non-native speakers at higher levels of attainment of the Croatian language both factors are influential. Furthermore, as opposed to non-native speakers, native speakers are insensitive to the morphological form of the collocation, which is expected considering that the participants are adult native speakers of Croatian for whom the morphological development is completed.

### Analysis of Answer Accuracy in the Task with Collocations in the Broader Context

With the aim of examining accuracy of answers in tasks of productive knowledge of collocations in the broader context a two-way analysis of variance with repeated measurements on factors of frequency (lower, higher) and associative strength (weak, strong) was carried out. The main effects of frequency,  $F(1, 69)=102.737, MSE=.036, p<.001$ , and associative strength  $F(1, 69)=116.197, MSE=.034, p<.001$ , were obtained in addition to the interaction of frequency and associative strength on the accuracy of answers,  $F(1, 69)=56.854, MSE=.037, p<.001$ . Duncan's post hoc test established that the accuracy of answers was greatest with collocations of high frequency and strong associative strength ( $M=.81; SE=.02$ ) in comparison with other types of collocations: collocations of higher frequency and weak associative strength ( $M=.40; SE=.03; p<.001$ ) and low frequency and strong ( $M=.42; SE=.03; p<.001$ ) or weak associative strength ( $M=.34; SE=.03; p<.001$ ) (Figure 2), which prove hypothesis H3.

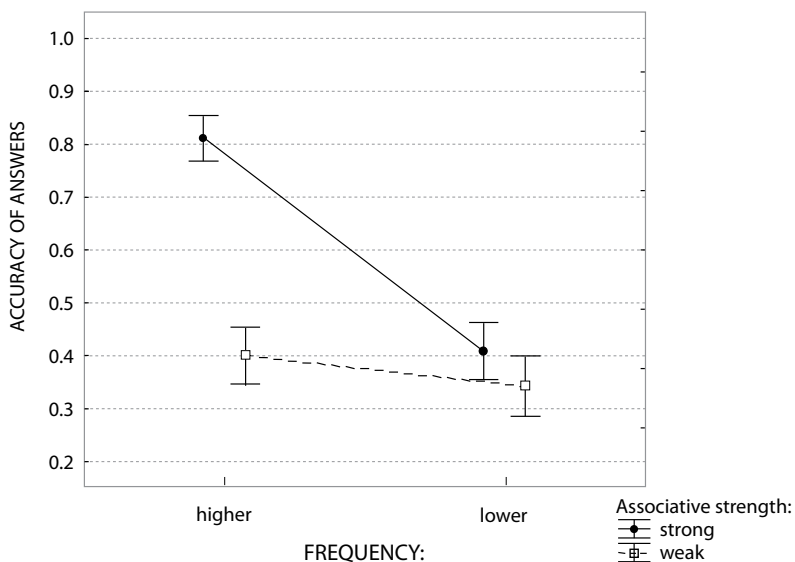


Figure 2. Accuracy of answers in tasks of productive knowledge of collocations in the broader context with respect to frequency (low, high) and associative strength (strong, weak) of collocations

### Analysis of Answer Types in Tasks at the Productive Level

An analysis of answers for all tasks was done in order to establish types of answers. Participants' answers were grouped into three categories of deviations (lexical, grammatical, and overlap). The categories and types of deviations are shown in Table 5. Answers including collocators out of context (KIK) and incorrect form of collocator (NOK) comprised those answers which were orthographically and semantically correct, and for category KIK grammatically correct.

Table 5  
Categories and types of deviations

| Category               | Type of answer  |
|------------------------|---|
| Lexical deviations     | Neologisms (NT) - lexical units emerging from word formation rules of the Croatian language                           |
|                        | Collocators out of context (KIK) – existing lexical combinations in the Croatian language, but out of the set context |
|                        | Unknown motivation of deviation (NM) – unclear criteria in collocator choice  |
| Grammatical deviations | Unknown motivation of deviation (NM) – unclear criteria in collocator choice  |
| Overlaps               | Collocators out of context + incorrect case form of collocator (KIK+NOK)  |

Following is the analysis which will explain and provide examples for all types of answers. The analysis of answer types in tasks of productive knowledge with collocations in the nominative (N=35) and oblique cases (N=35) and in the broader context encompassed 70 native Croatian speakers.

### ***Analysis of Answer Types in Tasks with Collocations in the Nominative and Oblique Cases***

Generally, it can be said that the average percent of correct answers is equal regardless of the task type, i.e. word form (Table 6). As for other types of answers, the most representative are collocators out of context, where the average percent is higher in the task with collocations in the nominative (34%) compared to the task with collocations in the oblique cases (27%). In the task with collocations in the nominative (Task A) the absolute percentage of correct answers (100%) was achieved for collocations *human rights* and *action movie* while the lowest percent was marked for collocations *status symbol* (11%) and *stereotype roles* (6%). In the task with collocations in the oblique cases (Task D) the absolute percentage (100%) was achieved for the collocation *action movie* while collocations *personal data* (94%) and *sports paper* (91%) had a somewhat smaller percentage. The lowest percentage of correct answers in the task with collocations in the oblique cases was achieved for collocations *nongovernment organizations* (26%) and *stereotype roles* (6%).

Table 6  
Average percent of a particular category of answer in the task with collocations in the nominative (A) and oblique cases (D)

|                | nominative | oblique cases |
|----------------|------------|---------------|
| CORRECT        | 61         | 61            |
| KIK            | 34         | 27            |
| TYPE OF ANSWER |            |               |
| KIK+NOK        | 0          | 2             |
| NT             | 1          | 0             |
| INCORRECT      | 3          | 9             |

In order to get a clearer perception of the type of answers of noun collocations among native Croatian speakers, an analysis of collocations of differing frequency and associative strength was carried out (Table 7).

Table 7  
 Percentage of answers with respect to frequency and associative strength

|                |            | nominative          |     |                 |     | oblique case     |     |                 |     |
|----------------|------------|---------------------|-----|-----------------|-----|------------------|-----|-----------------|-----|
|                |            | type of collocation |     |                 |     |                  |     |                 |     |
|                |            | higher frequency    |     | lower frequency |     | higher frequency |     | lower frequency |     |
|                |            | SAS                 | WAS | SAS             | WAS | SAS              | WAS | SAS             | WAS |
| TYPE OF ANSWER | CORRECT%   | 75                  | 67  | 63              | 40  | 74               | 68  | 66              | 36  |
|                | KIK%       | 22                  | 29  | 34              | 52  | 22               | 21  | 25              | 42  |
|                | NT%        | 1                   | 0   | 1               | 1   | 0                | 0   | 1               | 0   |
|                | KIK+NOK%   | 1                   | 0   | 0               | 1   | 1                | 2   | 3               | 3   |
|                | INCORRECT% | 1                   | 5   | 1               | 6   | 3                | 9   | 6               | 19  |

As expected, the lowest percentage of correct answers was recorded for collocations of lower frequency and weak associative strength. The highest percentage of incorrect answers was recorded for collocations of lower frequency and weaker associative strength in the task with morphologically marked collocations in the oblique cases. As for other types of answers, a relatively high percent was recorded for collocators out of context (KIK), particularly for collocations of lower frequency and weak associative strength. For instance, participants offered the following answers for particular collocations:

~hand bag: \*travel, \*passenger, \*smaller, \*small, \*portable, \*own, \*personal, \*big, \*sports, \*woman’s

~stereotype roles: \*business, \*everyday, \*main, \*house, \*family, \*daily, \*discriminatory, \*marital, \*social, \*assigned, \*film, \*sex, \*gender.

The analysis indicates that the percentage of correct answers is almost equal for collocations of higher frequency and strong or weak associative strength and lower frequency and strong associative strength. As opposed to non-native speakers (Ordulj, 2016), native speakers have an exceptionally smaller number of incorrect and grammatically inaccurate answers, which had been expected. However, it was not expected that native speakers, similar to non-native speakers, produced a relatively high number of answers of the type “collocation out of context”. The mentioned results indicate a need for a more detailed research of the role of context in attaining collocations.

### **Analysis of Answer Types in the Task with Collocations in the Broader Context**

The average percent of a particular type of answer shows an almost equal percentage of correct answers (49%) and collocators out of context (42%) (Table 8). The highest

percentage was recorded for collocations *human rights* (97%) and *life standard* (86%) while the lowest percentage was observed for collocations *purchase power* (16%) and *deadly diseases* (7%).

Table 8

*Average percentage of a particular category of answer in the task with collocations in the broader context*

|                | broader context | %  |
|----------------|-----------------|----|
|                | CORRECT         | 49 |
| TYPE OF ANSWER | KIK             | 42 |
|                | KIK+NOK         | 1  |
|                | INCORRECT       | 8  |

An in-depth analysis with respect to frequency and associative strength of collocations in the broader context showed a high percentage of correct answers among native speakers for collocations of high frequency and strong associative strength (Table 9).

Table 9

*Percentage of answers with respect to frequency and associative strength of collocations in tasks with collocations in the broader context*

|                |           | broader context     |     |                 |     |
|----------------|-----------|---------------------|-----|-----------------|-----|
|                |           | type of collocation |     |                 |     |
|                |           | higher frequency    |     | lower frequency |     |
|                |           | JAS                 | SAS | JAS             | SAS |
| TYPE OF ANSWER | CORRECT%  | 81                  | 40  | 41              | 34  |
|                | KIK%      | 16                  | 48  | 47              | 57  |
|                | NM%       | 0                   | 0   | 2               | 0   |
|                | KIK+NOK%  | 0                   | 2   | 0               | 0   |
|                | NT%       | 1                   | 0   | 0               | 0   |
|                | INCORRECT | 2                   | 11  | 10              | 9   |

For other types of collocations (HF and WAS, and LF and SAS or WAS) there was an observed high percentage of answers - collocators out of context (KIK). Correct answers for collocations of higher frequency and weak associative strength and low frequency and strong associative strength are approximately the same in number as well as for collocators out of context. Some examples of collocators out of context (KIS) are as follows:

*~fatal diseases*: \*malignant, \*remediable, \*incurable, \*dangerous, \*autoimmune, \*severe \*chronic, \*other, \*dangerous, \*cancerous, \*immune, \*frequent, \*benign, \*transmitted, \*deadly

*~tourist centers*: \*shopping, \*big, \*best, \*bigger, \*large, \*Dalmatian, \*city, \*market, \*main, \*urban, \*regional.



The quantitative analysis of the use of collocations in the broader context indicates that the broader context does not have a greater influence on the accuracy of answers, as their percentage is comparable with those in the narrower context (isolated sentences). However, as opposed to the use of collocations in isolated sentences, the use of collocations in the broader context showed greater qualitative differences with respect to factors of frequency and associative strength. The percentage of correct answers is greatest for collocations of high frequency and strong associative strength, and significantly lower, only one third, for collocations of low frequency and weak associative strength. This result leads to the conclusion that in research on acquisition of collocations, the relationship between context and different types of collocations should be examined.

## **Conclusion**

Research studies on collocations in the Croatian language to date are mostly limited to analyses of collocations as linguistic manifestations, while research on their acquisition in the Croatian as a mother tongue is lacking. The first step in researching attainment of collocations in the mother tongue was to examine factors that affect the use of collocations among native Croatian speakers. The research results point to several conclusions. Firstly, in using collocations in the mother tongue, factors of frequency of collocation and strength of associative strength are significant, while formal factors (type of collocation) are not important for native speakers. Differences in accuracy of use of collocations of varied frequency and associative strength are confirmed by the postulates of applied theories of language acquisition according to which varied types of frequency influence the acquisition of language items and patterns. Secondly, according to the observed use of collocations it is possible to compare native and non-native speakers. Native speakers are similar to non-native speakers at a high level of knowledge, however, significant differences between them are observed. As mentioned previously, native speakers find the factor of morphological form of the collocation insignificant as they, as native speakers, have acquired the morphological aspect of language. Thirdly, the influence of context in which a collocation is used is generally relatively small, however, if collocations of different characteristics are observed, the context is indicated as an important factor for the use of collocations of low frequency and weak associative strength, which is a very important finding for language teaching. Considering that the research revealed a relatively high percentage of produced collocations which do not match the set context (although they are certified collocations), it would be of importance to research (and describe) the relationship of collocation and context in which it occurs. The results of the research conducted are significant as they not only provide insight into the general processes of attainment of Croatian as the mother tongue and its use, but also provide a starting point for considering new methods and more efficient teaching approaches in the system of education.

## References

- Anić, V. (2003). *Veliki rječnik hrvatskoga jezika*. Zagreb: Novi liber.
- Bahns, J., & Eldaw, M. (1993). Should we teach EFL students collocations? *System*, 21(1), 104-114. [https://doi.org/10.1016/0346-251X\(93\)90010-E](https://doi.org/10.1016/0346-251X(93)90010-E)
- Begagić, M. (2014). English language students' productive and receptive knowledge of collocations. *Explorations in English Language and Linguistics*, 2(1), 46-67. <https://doi.org/10.1515/exell-2016-0003>
- Blagus Bartolec, G. (2014). *Riječi i njihovi susjedi. Kolokacijske sveze u hrvatskom jeziku*. Zagreb: Institut za hrvatski jezik i jezikoslovlje.
- Bonci, A. (2002). Collocational restrictions in Italian as a second language: a case control study. *Tuttitalia*, 26, 3-14.
- Borić, N. (2004). Kolokacije kao dio leksičkog pristupa u nastavi stranih jezika. In N. Ivanetić, B. Pritchard, & D. Stolac (Eds.), *Suvremena kretanja u nastavi jezika*, Zbornik HDPL (pp. 63 – 68). Zagreb-Rijeka: Graftrade.
- Borić, N. (2002). Leksikografska prezentacija kolokacijskih sklopova. *Strani jezici*, 31(1-2), 55-63.
- Borić, N. (1998). Semantički aspekt kolokacijskih odnosa s kontrastivnog stajališta. *Strani jezici*, 27(2), 72-79.
- Borić, N. (1996). *Imeničke kolokacije u hrvatskom i njihovi kontrastivni korespondenti i prijevodni ekvivalenti u engleskom jeziku. (Master's thesis)*. Zagreb: Filozofski fakultet.
- Burić, H., & Lasić, J. (2012). Kolokacije u nastavi inojezičnoga hrvatskog. In R. Relja (Ed.), *Zbornik radova Filozofskoga fakulteta u Splitu*, 5 (pp. 233-249). Split: Filozofski fakultet.
- Cvikić, L. (2016). *Hrvatski kao ini jezik: odabrane teme*. Zagreb: Hrvatsko filološko društvo.
- Cruse, D.A. (1986). *Lexical Semantics*. Cambridge: Cambridge University Press.
- Duplančić Rogošić, G. (2007). *Obrada kolokacija u englesko-hrvatskim, hrvatsko-engleskim i englesko-engleskim rječnicima. (Master's thesis)*. Zadar: Filozofski fakultet.
- Firth, J.R. (1957). *Modes of Meaning. Papers in Linguistics 1934-1951*. Oxford: Oxford University Press.
- Gitsaki, C. (1999). *Second Language Lexical Acquisition. A study of the Development of Collocational Knowledge*. San Francisco: International Scholars Publications.
- Greenbaum, S. (1974). Some verb-intensifier collocations in American and British English. *American Speech*, 49, 79-89. <https://doi.org/10.2307/3087920>
- Gulešić Machata, M., & Machata, M. (2007). Boje u hrvatskim i slovačkim kolokacijama. *Riječ*, 13(2), 99-107.
- Halliday, M.A.K. (1966). Lexis as a linguistic level. In C.E. Bazell, J.C. Catford, M.A.K. Halliday, & R.H. Robins (Eds.), *In memory of J.R. Firth* (pp. 148-162). London: Longman.
- Henriksen, B. (2013). Research on L2 learners' collocational competence and development – a progress report. In C.B. Bardel, B. Laufer, & C. Lindqvist (Eds.), *L2 vocabulary acquisition, knowledge and use* (pp. 29-56). Eurosla Monographs Series, 2. Eurosla. Retrieved from <http://www.eurosla.org/monographs/EM02/EM02tot.pdf>
- hrWaC – Croatian web corpus. Retrieved from <http://nlp.ffzg.hr/resources/corpora/hrwac/>

- Jaén, M.M. (2007). A Corpus-driven Design of a Test for Assessing the ESL Collocational Competence of University Students. *International Journal of English Studies*, 7(2), 127-147.
- Katz, J.J., & Fodor, J.A. (1963). The structure of a semantic theory. *Language*, 39(2), 170-210. <https://doi.org/10.2307/411200>
- Kos, I. M. (2012). *Kolokacijske strukture u hrvatskome naftnom nazivlju. (Master's thesis)*. Zagreb: Filozofski fakultet.
- Koya, T. (2003) A study of Collocation in English and Japanese Noun-Verb Combinations. *Intercultural Communication Studies XII.*, 125-145.
- Lewis, M. (2000). *Teaching collocation. Further developments in the Lexical Approach*. England: Language Teaching Publications.
- Ljubešić, N., & Erjavec, T. (2011). hrWaC and slWac: Compiling Web Corpora for Croatian and Slovene. In I. Habernal, & V. Matoušek (Eds.), *Text, Speech and Dialogue. TDS 2011. Lecture Notes in Computer Science*, vol. 6836 (pp. 395-402). Berlin – Heidelberg: Springer.
- Martyńska, M. (2004). Do English language learners know collocations? *Investigationes Linguisticae*, 12, 1-12. <https://doi.org/10.14746/il.2004.11.4>
- Matthews, P.H. (1965). Problems of selection in transformational grammar. *Journal of Linguistics*, 1, 35-47. <https://doi.org/10.1017/S0022226700001006>
- Melčuk, I. (2001). *Collocations and Lexical Functions*. In A. P. Cowie (Eds.), *Phraseology Theory, Analysis and Applications* (pp. 23-53). Oxford-New York: Oxford University Press.
- Mihaljević, M. (1991). Višerječne natuknice i podnatuknice u jednojezičnom općem rječniku hrvatskoga jezika. *Rasprave Instituta za hrvatski jezik i jezikoslovlje*, 17, 133–143.
- Mikelić Preradović, N., Berać, M., & Boras, D. (2015). Learner Corpus of Croatian as a Second and Foreign Language. In K. Cergol Kovačević, & S. L. Udier (Eds.), *Multidisciplinary Approaches to Multilingualism* (pp. 107-126). Frankfurt am Main, Germany: Peter Lang,
- Miščin, E. (2012). *Glagolske kolokacije u engleskome jeziku medicinske struke. (Doctoral thesis)*. Osijek: Filozofski fakultet.
- Miščin, E., & Pavičić Takač, V. (2013). Analiza kolokacijske kompetencije neizvornih korisnika engleskog jezika medicine. *JAH*, 4(7), 235-256.
- Mitchell, T.F. (1971). Linguistic goings on: collocations and the other lexical matters arising on the syntagmatic record. *Archivum Linguisticum*, 2, 35-61.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. United Kingdom: Cambridge University Press. <https://doi.org/10.1017/CBO9781139524759>
- Ordulj, A. (2016). *Kolokacije kao sastavnica leksičke kompetencije u hrvatskome kao inome jeziku. (Doctoral thesis)*. Zagreb: Filozofski fakultet.
- Pavičić Takač, V., & Lukač, M. (2013). How word choice matters: An analysis of adjective-noun collocations in a corpus of learner essays. *Jezikoslovlje*, 14(2-3), 385-402.
- Petrović, B. (2008a). Nacrtak za Kolokacijski rječnik hrvatskoga jezika. In M. Samardžija (Ed.), *Vidjeti Ohrid: Referati hrvatskih sudionika i sudionika za XIV. Međunarodni slavistički kongres* (pp. 225–235). Zagreb: Hrvatsko filološko društvo – Hrvatska sveučilišna naklada.
- Petrović, B. (2008b). Glagoli emocionalnih stanja u kolokacijskim strukturama i leksikografskom opisu. In I. Srdoč-Konestra, & S. Vranić (Eds.), *Riječki filološki dani 7:*

- Zbornik radova s Međunarodnoga znanstvenog skupa Riječki filološki dani održanog u Rijeci od 16. do 18. studenoga 2006.* (pp. 589–599). Rijeka: Filozofski fakultet.
- Petrović, B. (2007). Razvijanje kolokacijske kompetencije u hrvatskome kao stranom i drugom jeziku. *Strani jezici*, 36, 31-38.
- Prichard, B. (1998). O kolokacijskom potencijalu rječničkog korpusa. *Filologija*, 30-31, 285-299.
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Sinclair, J. (1966). Beginning the study of lexis. In C. E. Bazell, J. C. Catford, M.A.K. Halliday, & R. H. Robins (Eds.), *In memory of J.R. Firth* (pp. 410-430). London: Longman.
- Syanova-Chanturia, A., & Spina, S. (2015). Investigation of Native Speaker and Second Language Learner Intuition of Collocation Frequency. *Language Learning*, 65(3), 533-562. <https://doi.org/10.1111/lang.12125>
- Syanova, A., & Schmitt, N. (2008). L2 learner production and processing of collocation: A multi-study perspective. *Canadian Modern Language Review*, 64, 429-458. <https://doi.org/10.3138/cmlr.64.3.429>
- Špiranec, I. (2005). Priroda i upotreba kolokacija: primjeri iz tehničkoga engleskog jezika. *Strani jezici*, 34(3), 219-227.
- Štefić, L., Mravak–Stipetić, M., & Borić, V. (2010). Kolokacije u jeziku stomatologije: primjeri iz oralne medicine. *Acta stomatologica Croatica*, 44(3), 176-187.
- Turk, M. (2010). Status kolokacija u hrvatskim jednojezičnim i dvojezičnim rječnicima. In D. Bačić-Karković et al. (Eds.), *Riječki filološki dani: Zbornik radova sa 6. međunarodnoga znanstvenog skupa Riječki filološki dani* (pp. 537–547). Rijeka: Filozofski fakultet.
- Turk, M. (2000). Višečlani izrazi s desenantiziranom sastavnicom kao nominacijske jedinice. In D. Stolac (Ed.), *Riječki filološki dani. Zbornik radova 3* (pp. 477-486). Rijeka: Filozofski fakultet.

---

**Antonia Ordulj**

Croaticum – Center for Croatian as a Second and Foreign Language, Faculty of Humanities and Social Sciences, University of Zagreb  
Ivana Lučića 3, 10000 Zagreb, Croatia  
[antoniasvetic@gmail.com](mailto:antoniasvetic@gmail.com)

**Lidija Cvikić**

Faculty of Teacher Education, University of Zagreb, Savska 77, 10000 Zagreb, Croatia  
[lidija.cvikić@ufzg.hr](mailto:lidija.cvikić@ufzg.hr)

# Čimbenici ovladanosti kolokacijama kod izvornih govornika hrvatskoga jezika

---

## Sažetak

Istraživanja kolokacija u hrvatskome jeziku uglavnom su usmjerena na njihov lingvistički opis i leksikografsku problematiku, a istraživanja njihove uporabe rijetka su i uglavnom usmjerena na inojezične govornike (Burić i Lasić, 2012; Ordulj, 2016; Petrović, 2007). U ovome se radu donose rezultati istraživanja uporabe kolokacija kod izvornih govornika hrvatskoga jezika. Na temelju obavljenih korpusnih istraživanja obilježja kolokacija i ispitivanja govornika hrvatskoga kao materinskoga jezika, istražen je utjecaj dvaju čimbenika na uporabu kolokacija: čestote kolokacija te asocijativne snage kolokacijskih sastavnica. Istražena je uporaba kolokacija u morfološki neobilježenome obliku (nominativu), u kosim padežima, kao i uporaba u užem i širem kontekstu. Rezultati istraživanja pokazuju da su u uporabi kolokacija kod izvornih govornika ključni čimbenici čestote i asocijativne snage, a da je morfološki oblik kolokacije izvornim govornicima nebitan. Istaživanje je također pokazalo veći utjecaj konteksta na uporabu kolokacija manje čestote i slabije asocijativne snage od uporabe kolokacija veće čestote i jake asocijativne snage. Dobiveni rezultati važni su za razumijevanje općih procesa ovladavanja hrvatskim kao materinskim jezikom i kao polazište za promišljanja o učinkovitijim načinima poučavanja hrvatskoga jezika.

**Ključne riječi:** asocijativna snaga kolokacije; čestota; kontekst; ovladavanje materinskim jezikom.

## Uvod

Brojna lingvistička istraživanja već od ranih pedesetih godina prošloga stoljeća naglašavaju kako su kolokacije zanemarena sastavnica leksičke kompetencije inoga jezika (Gitsaki, 1999; Petrović, 2007), no u posljednjih se dvadesetak godina spomenutom leksičkom fenomenu posvećuje sve više pozornosti (Bahns i Eldaw, 1993; Gitsaki, 1999; Jaén, 2007; Koya, 2003; Martyńska, 2004). Radovi su koji se bave kolokacijskom kompetencijom u materinskom jeziku sporadični ponajprije stoga što je uvriježeno mišljenje kako izvorni govornici zbog neprestane izloženosti i uporabe materinskoga jezika intuitivno kombiniraju leksičke jedinice (usp. Borić, 2002). Iako

kolokacije i u hrvatskome jezikoslovlju sve više izazivaju pažnju, ponajprije za engleski kao ini jezik, za hrvatski su jezik radovi još uvijek sporadični. Dosad najiscrpniji opis imenskih kolokacija u hrvatskom kao materinskom jeziku (J1) donosi Blagus Bartolec (2014), a kolokacijama u hrvatskom kao inom jeziku (HIJ) bave se Petrović (2007), Burić i Lasić (2012), a Ordulj (2016) po prvi put donosi iscrpnu analizu kolokacija kao sastavnice leksičke kompetencije u hrvatskome kao inome jeziku.

Termin kolokacija najčešće se definira kao supojavljivanje pojedinačnih leksičkih jedinica na sintagmatskoj razini (Borić, 2004; Firth, 1957; Lewis, 2000; Nation, 2000; Petrović, 2007). Na taj se način značenje pojedinačnih leksičkih jedinica udružuje radi ostvarenja cjelovitoga značenja. Sam je pojam u teorijsku lingvistiku uveo John R. Firth (1957) koji o kolokacijama promišlja isključivo iz perspektive leksičkih odnosa i leksika kao samostalne razine proučavanja, odvojene od gramatike. Kolokacije je potrebno razlikovati od drugih sličnih sveza riječi: frazema, onima i naziva. Turk (2000) upozorava na sličnosti i razlike među njima. Među sličnosti Turk (2000) ubraja složenu strukturu, semantičku pretvorbu barem jedne sastavnice i reproduktivnost dok se stabilnost, konotativnost, idiomatičnost te semantički odnosi ubrajaju u razlike. Važno je obilježje u razlikovanju spomenutih leksičkih komunikacijska funkcija. Frazemi se smatraju izrazima sekundarne nominacije te ukazuju na osobne, emotivne i prosudbene stavove, nazivi služe za sporazumijevanje i definiranje pojmova među stručnjacima nekog znanstvenoga ili stručnoga područja, a onimima je primarna zadaća identificirati neki objekt. Blagus Bartolec (2014, str. 80) navodi kako je kolokacija „posebna leksička sveza na sintagmatskoj razini temeljena na značenjskoj povezanosti samostalnih leksičkih sastavnica u kojoj se konkretiziraju njihova značenja.” Blagus Bartolec (2014) kolokaciju ograničava na dvije punoznačne riječi koje na formalnom planu odgovaraju sintagmi.

Istraživanja kolokacija mogu se klasificirati s obzirom na različite pristupe proučavanja kolokacija na: leksičko-kompozicijski (Firth, 1957; Halliday, 1966; Sinclair, 1966), semantičko-sintaktički (Cruse, 1986; Katz i Fodor, 1963; Melčuk, 2008) i strukturalni pristup (Greenbaum, 1974; Matthews, 1964; Mitchell, 1971). Predstavnicima leksičko-kompozicijskoga pristupa leksik smatraju autonomnim entitetom i neovisnim o gramatici. Firth (1957) po prvi put uvodi pojmove kolokabilnosti i kolokacijskoga značenja te pokušava objasniti sam pojam kolokacije. Kolokaciju definira kao način realizacije leksičkoga značenja koje proizlazi iz uobičajenoga supojavljivanja dva leksema uzduž sintagmatske osi (Firth, 1957). Firthova teorijska promišljanja razrađuju njegovi učenici Halliday (1966) i Sinclair (1966). Iako ne nude odgovore na temeljna pitanja kolokacijske problematike, Firth (1957), Halliday (1966) i Sinclair (1966) zaslužni su za uvođenje pojma kolokacije, rad na terminologiji i za analizu značenjskih odnosa među leksičkim jedinicama na sintagmatskoj razini. Semantičko-sintaktičkim pristupom (Cruse, 1986; Katz i Fodor, 1963; Melčuk, 2008) pokušava se na temelju semantičkih obilježja predvidjeti kolokacijski raspon pojedine leksičke jedinice. Ni taj pristup ne uspijeva objasniti značenjske odnose među kolokacijskim

sastavnicama pa se veći naglasak stavlja na proučavanje paradigmatičkih leksičkih odnosa ostavljajući kolokacije po strani i razumijevajući ih kao iznimke od općega pravila. Strukturalnim pristupom zagovara se jedinstvo leksika i gramatike, a sve s ciljem utvrđivanja kolokacijske strukture. Istraživanja predstavnika strukturalnoga pristupa (Greenbaum, 1974; Matthews, 1964; Mitchell, 1971) pokušaj su da se kolokacije objasne uz pomoć strukturalnih uzoraka uzimajući u obzir i kolokacijska ograničenja. Unatoč brojnosti istraživanja i teorijskih pristupa, još uvijek nema odgovora na temeljna pitanja kolokacijske problematike odnosno zbog čega riječi kolociraju i na temelju kojih kriterija, kako jasno razlikovati kolokacije od drugih leksičkih sveza i kako nastaje kolokacijsko značenje.

Kolokacijama se u hrvatskoj literaturi pristupa s različitih aspekata te se, stoga, može govoriti o radovima koji se bave kontrastivnom, semantičkom i leksikografskom obradom kolokacija te kolokacijama u jeziku struke. Također, primjetan je veći interes za glotodidaktički aspekt proučavanja kolokacija koji naglasak stavlja na razvoj kolokacijske kompetencije pri čemu u prvi plan dolazi učenje i poučavanje leksičkih kombinacija koje se supojavljaju na sintagmatskoj razini.

Radovi koji se bave kontrastivnom i semantičkom obradom kolokacija (Borić, 1998; 1996; Gulešić Machata & Machata, 2007) u prvom redu govore o semantičkim odnosima među kolokacijskim sastavnicama te nastanku kolokacijskoga značenja, a obrađuje se i pitanje prijevodne ekvivalencije kolokacija. U okviru semantičkih rasprava izdvaja se rad Blagus Bartolec (2014) s obzirom na to da po prvi put pokušava definirati i opisati imenske kolokacije u hrvatskom jeziku te nastanak kolokacijskoga značenja postupkom sintagmatizacije. Naime, kolokacije u hrvatskome jeziku nastaju postupkom sintagmatizacije, što znači da nemaju status leksičkih jedinica jer dolazi do značenjske prilagodbe pod utjecajem značenjskoga potencijala druge riječi koja ulazi u kolokacijski odnos. Zahvaljujući značenjskom potencijalu kolokacijskih sastavnica, dolazi do nadopunjavanja jedne sastavnice drugom i oblikovanja novoga, kolokacijskoga značenja koje je „motivirano izvanjezičnom pojavom na koju referira (npr. *sunčan/kišan dan*, *duga/kratka kosa*) (Blagus Bartolec, 2014 str. 85).” Blagus Bartolec (2014) također po prvi put donosi i podjelu kolokacija prema vrstama s obzirom na formalna i značenjska obilježja leksičkih jedinica koje se supojavljaju. Brojčano su najzastupljeniji radovi koji se bave leksikografskom obradom kolokacija u hrvatskim rječnicima (Borić, 2002; Duplančić Rogošić, 2007; Mihaljević, 1991; Petrović, 2008a, 2008b; Pritchard, 1998; Turk, 2010) s obzirom na to da je još uvijek otvoreno pitanje uvrštavanja kolokacija u rječnike te s obzirom na činjenicu da za hrvatski još uvijek ne postoji kolokacijski rječnik. Zanimanje za kolokacije u jeziku struke, ponajprije engleskom, pokazuje sve više autora nastojeći ukazati na problematiku s obzirom na posebnost rječnika pojedine struke. Najviše je radova o kolokacijama u jeziku medicinske (Štefić, Mravak–Stipetić & Borić, 2010; Miščin, 2012; Miščin & Pavičić Takač, 2013) i tehničke struke (Kos, 2012; Špiranec, 2005).

Kao što je navedeno, uz iznimku nekolicine radova o hrvatskome kao inome jeziku (Burić i Lasić, 2012; Ordulj, 2016), u hrvatskome jezikoslovlju općenito nedostaju



istraživanja o razvoju kolokacijske kompetencije. Ponajprije je važno istražiti razvoj kolokacijske kompetencije u hrvatskome kao materinskome jeziku jer uporaba kolokacija može biti važan pokazatelj stupnja leksičkoga znanja. Uz to, tek se opisom kolokacijske kompetencije u hrvatskome kao materinskome jeziku može definirati ciljana razina ovladavanja hrvatskim kao inim jezikom. Stoga je osnovni cilj ovoga rada istražiti čimbenike koji utječu na kolokacijsku kompetenciju izvornih govornika hrvatskoga jezika kako bi se dobila jasnija slika ovladavanja kolokacijama u hrvatskome.

### ***Kolokacijska kompetencija***

Kolokacijska je kompetencija jedan od najvažnijih segmenata produktivnoga i receptivnoga leksičkoga znanja s obzirom na to da kolokacije i izvornim i neizvornim govornicima omogućuju da razlikuju značenjske varijante polisemnih riječi i da razumiju konotativno značenje (Henriksen, 2013). Također, kolokacije pridonose fluentnosti i prirodnijoj komunikaciji te sužavaju značenje pojedinačnih leksičkih jedinica koje ulaze u kolokacijski odnos stvarajući asocijativne veze između kolokacijskih sastavnica (Borić, 2004; Henriksen, 2013; Nation, 2001). Na važnost kolokacija upućuju brojni primijenjeni lingvisti, ponajprije kada je riječ o ovladavanju inim jezikom, te naglašavaju kako se kolokacijama u nastavi inoga jezika ne posvećuje dovoljno pažnje te im treba pristupiti sustavnije i poučavati ih eksplicitnije (Bahns i Eldaw, 1993). Općenito se smatra da se s razvojem općega vokabulara razvija i kolokacijska kompetencija (Bonci, 2002; Gitsaki, 1999; Koya, 2003). Pritom učenici na višim jezičnim razinama pokazuju bolje znanje leksičkih i gramatičkih kolokacija u inom jeziku (Begagić, 2014; Bonci, 2002; Gitsaki, 1999). Na razvoj kolokacijske kompetencije utječe nekoliko čimbenika, a među najvažnijima su frekvencija izloženosti kolokacijama, razlike u materinskom jeziku, strukturna i semantička složenost kolokacijskih vrsta, ali i obrazovno okruženje, odnosno izloženost različitim kolokacijskim vrstama putem nastavnog procesa (Begagić, 2014; Gitsaki, 1999).

Osim radova koji se bave isključivo kolokacijskom kompetencijom u inom jeziku, noviji se radovi bave usporedbom procjene frekventnosti kolokacija kod izvornih i neizvornih govornika. Istraživanje Siyanova-Chanturie i Spine (2015) pokazuje da se izvorni i neizvorni govornici talijanskoga jezika ne razlikuju u procjeni visoko frekventnih kolokacija pa njihova procjena pozitivno korelira s frekvencijom kolokacija iz korpusa. U istome su istraživanju izvorni govornici bili uspješni u procjeni kolokacija jako niske frekventnosti, kao i naprednija skupina govornika talijanskoga kao inoga jezika, no najveći problem za obje skupine govornika činile su kolokacije srednje frekventnosti jer se one nalaze u sredini kontinuuma. Naime, u poučavanju inoga jezika pozornost se pridaje visokofrekventnim jezičnim jedinicama jer su česte, a time i korisne za komunikaciju, te niskofrekventnima, za koje je mala vjerojatnost da će im učenici jezika biti izloženi u jezičnome unosu, pa ih je potrebno eksplicitno poučavati. Siyanova i Schmitt (2008) nastojali su utvrditi koliko se neizvorni

govornici koriste prikladnim kolokacijama u pisanoj produkciji te jesu li uspješni u prosudbi stupnja čestote kolokacija u odnosu na izvorne govornike engleskoga jezika. Istraživanje je potvrdilo kako se neizvorni govornici koriste česte i asocijativno jakim kolokacijama u istoj mjeri kao i izvorni govornici.

Stoga je istraživanje koje se predstavlja u ovome radu usmjereno upravo na ta dva ključna čimbenika uporabe kolokacija: čestotu i asocijativnu jačinu kolokacijskih sastavnica.

## **Metodologija istraživanja**

Temeljni je cilj rada ispitati uporabno znanje imenskih kolokacija s obzirom na čestotu i asocijativnu jačinu kolokacijskih sastavnica u hrvatskome kao materinskome jeziku.

U radu su postavljeni sljedeći problemi:

1. Ispitati znanje imenskih kolokacija s obzirom na čestotu i jačinu kolokacijskih sastavnica.
2. Ispitati produktivno znanje morfološki neobilježenih kolokacija u nominativu i morfološki obilježenih kolokacija u kosim padežima s obzirom na jačinu asocijativne snage kolokacijskih sastavnica i čestotu kolokacija.
3. Ispitati produktivno znanje imenskih kolokacija u širem kontekstu.

S obzirom na istraživačka pitanja, postavljene su sljedeće hipoteze:

- H1: Ispitanici će bolje znati kolokacije više čestote i jače asocijativne snage od kolokacija niže čestote i slabije asocijativne snage.
- H2: Ispitanici će pokazati podjednaku točnost odgovora kod morfološki neobilježenih kolokacija u nominativu i morfološki obilježenih kolokacija u kosim padežima, neovisno o čestoti i asocijativnoj snazi kolokacija.
- H3: Kod kolokacija u širem kontekstu, ispitanici će imati podjednaku točnost odgovora kod kolokacija više čestote i jače asocijativne snage.

Ispitivanje se kolokacijske kompetencije u ovom radu odvijalo u dvije faze. U prvoj je fazi bilo potrebno provesti korpusno istraživanje kolokacija u korpusu hrWaC (Croatian Web Corpus, Ljubešić & Erjavec, 2011; hrWaC) te potom ispitati asocijativnu snagu kolokacijskih sastavnica. U drugoj je fazi izrađen instrument i provedeno je istraživanje na ispitanicima, izvornim govornicima hrvatskoga jezika.

## **Odabir kolokacija za istraživanje**

Za samo je istraživanje bilo potrebno prikupiti nekoliko stotina kolokacija koje su predstavljale radni korpus te su se koristile u sastavljanju instrumenata. U istraživanju prevladavaju kolokacije osnovnoga tipa u kojima se ostvaruje primarno značenje, odnosno kako Blagus Bartolec (2014, str. 93) navodi dolazi do „osnovnoga stupnja konkretizacije primarnoga značenja.” U korpusnoj analizi na kojoj se temeljio odabir

kolokacija prihvaćena je binarna podjela kolokacijskih sastavnica na osnovu i kolokat pri čemu je osnova bila imenica, a kolokat pridjev. Kolokacije su odabrane na temelju čestotnosti (viša i niža) te asocijativne snage kolokacijskih sastavnica (jake i slabe kolokacije).

Za istraživanje s izvornim govornicima odabran je isti ispitni materijal kao i za istraživanje Ordulj (2016) s neizvornim govornicima hrvatskoga jezika u kojemu su se kolokacije crpile iz eseja polaznika neizvornih govornika hrvatskoga jezika (iscrpnije u Ordulj, 2016). Ukupno je analizirano 300 eseja koji su dio prvoga Učeničkoga korpusa hrvatskoga kao inoga jezika – CROLTEC (Mikelić Preradović, Berać, & Boras, 2015) koji još uvijek nije javno dostupan.

Ovjeranost kolokacija provjerena je u Anićevu (2003) Velikom rječniku hrvatskoga jezika te potom u korpusu hrWaC (Ljubešić i Erjavec, 2011) koji trenutno predstavlja najveći korpus za hrvatski jezik s ukupno 1.9 milijardi pojava. Korpus hrWaC prvi je web korpus za hrvatski jezik koji obuhvaća sve mrežne stranice s nacionalne domene Republike Hrvatske (.hr). U ovom je radu za pretragu kolokacija u korpusu hrWaC korišten programski jezik CQL (eng. *Corpus Query Language*) koji omogućuje pretraživanje korpusa s označenim riječima s obzirom na različite vrste atributa (eng. *default attribute*). Za potrebe ovoga rada kolokacije su se pretraživale prema atributu *riječ*, a CQL upit za kolokaciju *osnovna škola* izgleda ovako: [word=„osnovn.\*” & tag=„A.\*”] [word=„škol.\*” & tag=„N.\*”]

Pomoću računalnojezičnih alata u korpusu hrWaC izračunala se i čestotnost prikupljenih kolokacija koje su prema čestoti podijeljene u dvije skupine – više (VČ) i niže (NČ) čestote. Takva je podjela u skladu sa Sinclair (1991) koji navodi da će čestota traženih pojmova ovisiti o vrsti i broju pojava unutar korpusa, a što je korpus veći, i rezultati će biti vjerodostojniji.

### ***Ispitanici***

Istraživanje je kolokacijske kompetencije u hrvatskom kao materinskome jeziku provedeno tijekom zimskoga semestra ak. god. 2015./2016. U istraživanju procjene asocijativne snage kolokacija sudjelovao je 50 studenata Filozofskoga fakulteta u Rijeci, a u istraživanju kolokacijske kompetencije 70 studenata Učiteljskoga fakulteta u Zagrebu (Ž=65, M=5). Ispitanici u oba istraživanja izvorni su govornici hrvatskoga jezika.

### ***Instrumenti***

U istraživanju su se koristila dva instrumenta izrađena za potebe istraživanja. Jednim se ispitala procjena asocijativne snage kolokacija, a drugim kolokacijska kompetencija izvornih govornika.

Asocijativna jačina kolokacijskih sastavnica procjenjivala na skali od 1 do 5, a kolokacije su bile raspoređene slučajnim redoslijedom. Za svaku su kolokaciju izračunati deskriptivni podatci (M, SD) koji predstavljaju prosječnu asocijativnu snagu.

Iako se radi o subjektivnim procjenama, uprosječni podaci služe kao objektivni pokazatelj asocijativne snage, odnosno jačine između kolokacijskih sastavnica. S obzirom na vrijednosti aritmetičke sredine (M) nakon istraživanja, kolokacije su grupirane u dvije skupine – jake (JAS) i slabe asocijativne snage (SAS).

Kolokacijska kompetencija ispitivala se instrumentom koji se sastojao od tri zadatka popunjavanja praznina: a) zadatak s kolokacijama u morfološki neobilježenom padežu nominativu; b) zadatak s kolokacijama u kosim padežima; c) zadatak s kolokacijama u širem kontekstu. Prvi je zadatak imao ukupno dvadeset rečenica u kojima su kolokacije bile u nominativu (zadatak A), a u drugom su zadatku iste kolokacije bile u kosim padežima (zadatak D) (Tablica 1). U svim su zadacima bile zastupljene kolokacije s obzirom na istraživana obilježja: čestotu i asocijativnu snagu. Odnosno, u zadacima su bile zastupljene kolokacije visoke čestote i jake asocijativne snage (VČ + JAS), visoke čestote i slabe asocijativne snage (VČ + SAS), niske čestote i jake asocijativne snage (NČ + JAS) te niske čestote i slabe asocijativne snage (NČ + SAS).

#### Tablica 1

U trećem su zadatku kolokacije bile uklopljene u informativniji, širi kontekst (zadatak C). U njemu je bilo zastupljeno deset kolokacija u kosim padežima iz D zadatka i deset novih kolokacija koje ovaj zadatak ne dijeli s prethodnim zadacima (Tablica 2).

#### Tablica 2

Istraživanjem je bio obuhvaćen ravnomjeran broj kolokacija s obzirom na kriterij čestote i jačine asocijativne snage kolokacijskih sastavnica. U svim je zadacima bila zadana osnova kolokacije (imenica), a ispitanici su trebali pokazati produktivno znanje kolokata (pridjev). Kao točni odgovori uzimali su se u obzir oni kolokati koje su predvidjele autorice testova, odnosno ovoga istraživanja. Svi ostali odgovori s obzirom na semantičke, tvorbene, gramatičke ili pravopisne sličnosti ili razlike, grupirani su u različite vrste odgovora u odabiru kolokata ciljane kolokacije.

Rezultati su se obrađivali kvantitativno i kvalitativno (tipologija vrsta odgovora u odabiru kolokata). Koristio se statistički paket Statistica.

### ***Postupak provedbe istraživanja***

Ispitanici u istraživanju bili su upoznati sa svrhom istraživanja te je njihovo sudjelovanje bilo dobrovoljno i anonimno. S obzirom na specifičnost ispitnoga materijala, odnosno činjenicu da su u zadatku A (nominativ) i zadatku D (kosi padeži) bile zastupljene iste kolokacije te je i zadatak C (kolokacije u širem kontekstu) dijelio deset istih kolokacija sa zadatkom D, provedba je istraživanja unutar svake skupine organizirana prema unaprijed određenoj shemi, a redosljed kolokacija u svim zadacima variran je te je svaki ispitanik imao lozinku. Ispitni su se materijali dijelili pojedinačno.

Kako bi se ispitao utjecaj morfoloških obilježja riječi na znanje kolokacija, ispitanici su unutar skupina podijeljeni u dvije ravnomjerne podskupine. Prva je skupina rješavala zadatak A s kolokacijama u nominativu, a druga zadatak D s kolokacijama u kosim padežima. Raspodjela se zadataka unutar svake skupine odvijala prema rasporedu prikazanom u Tablici 3.

Tablica 3

Sljedeća poglavlja donose analizu rezultata istraživanja kolokacijske kompetencije u hrvatskom kao materinskom jeziku (J1)<sup>1</sup>. Najprije se donose opći rezultati, a potom specifični s obzirom na obilježja kolokacija (VČ, NČ, JAS, SAS) i zadatke. Također, zbog preglednosti i boljega razumijevanja, interpretacija se rezultata donosi odmah nakon analize podataka.

## Rezultati i rasprava

Najprije će biti predstavljena analiza točnosti odgovora u zadacima produktivne razine, najprije u zadacima s kolokacijama u nominativu, zatim s kolokacijama u kosim padežima i naposljetku produktivno znanje kolokacija u širem kontekstu.

### ***Analiza točnosti odgovora u zadacima s kolokacijama u nominativu i kosim padežima***

S ciljem ispitivanja točnosti odgovora u zadacima produktivnoga znanja morfološki neobilježenih kolokacija u nominativu (Zadatak A) i morfološki obilježenih kolokacija u kosim padežima (Zadatak D) provedena je trosmjerna analiza varijance s ponovljenim mjerenjima na faktorima čestote (viša, niža) i asocijativne snage (jaka, slaba) kolokacija te nezavisnim skupinama na faktoru oblik riječi (nominativ, kosi padež). Rezultati su prikazani u Tablici 4.

Tablica 4

Dobiveni su glavni efekti čestote i asocijativne snage kolokacija te značajna dvosmjerna interakcija čestote i asocijativne snage kolokacija na točnost odgovora u zadacima produktivnoga znanja kod izvornih govornika hrvatskoga jezika. Duncanovim post hoc testom utvrđeno je da je točnost odgovora najviša kod kolokacija više čestote i jake asocijativne snage ( $M = ,75$ ;  $SE = ,02$ ) te da se ona značajno razlikuje u odnosu na sve preostale kolokacije (sve  $p < ,01$ ). Nadalje, podjednaka razina točnosti dobivena je kod kolokacija više čestote i slabe asocijativne snage ( $M = ,67$ ;  $SE = ,02$ ) te niže čestote i jake asocijativne snage ( $M = ,65$ ;  $SE = ,02$ ). Točnost odgovora najniža je kod kolokacija niže čestote i slabe asocijativne snage ( $M = ,38$ ;  $SE = ,03$ ).

<sup>1</sup> U provedenom se istraživanju jedinim zadatkom provjeravalo receptivno znanje kolokacija s ciljem provjere prikladnosti instrumenta za ispitivanje neizvornih govornika. S obzirom na to da je točnost odgovora bila iznimno visoka (za kolokacije više čestote više čestote  $M = ,99$ ;  $SD = ,03$ , a za kolokacije niže čestote  $M = ,97$ ;  $SD = ,06$ ), ti se odgovori neće analizirati u ovome radu. Ipak je važno naglasiti da se i u tome slučaju pokazala značajna razlika u točnosti odgovora  $t(69) = 2.189$ ,  $p < ,05$  s obzirom na čestotu kolokacija.

Dobivena razlika također je značajna u usporedbi sa svim preostalim kolokacijama (sve  $p < ,001$ ) (Slika 1). Navedeni rezultati potvrđuju hipoteze H1 i H2.

#### Slika 1

Dobiveni rezultati pokazali su da su izvorni govornici proizveli najtočnije odgovore za kolokacije visoke čestote i jake asocijativne snage. Taj se rezultat može dovesti u vezu s rezultatom istraživanja Siyanova i Schmitt (2008) koje je pokazalo da se neizvorni govornici najviše približavaju izvornima upravo u uporabi kolokacija visoke čestote i jake asocijativne snage. Takvi su rezultati u skladu sa spoznajama uporabno utemeljenih teorija ovladavanja jezikom (vidi Cvikić, 2016) prema kojima su različite vrste čestote (npr. čestota natuknice, čestota paradigme i slično) važni čimbenici ovladavanja pojedinim jezičnim jedinicama ili obrascima. Stoga bi u daljnjim istraživanjima valjalo ispitati kako na usvajanje kolokacije utječe čestota same kolokacije u odnosu na čestotu njezinih sastavnica. Dobiveni su rezultati pokazali i bitnu razliku u razvoju kolokacijske kompetencije izvornih i neizvornih govornika hrvatskoga. Naime, istraživanje Ordulj (2016) pokazalo je da kod neizvornih govornika na nižim razinama ovladanosti hrvatskim jezikom na proizvodnju kolokacija utječe samo čimbenik čestote, ali ne i asocijativne snage, dok na proizvodnju kolokacija kod neizvornih govornika na višim razinama ovladanosti hrvatskim jezikom utječu oba čimbenika. Uz to, za razliku od neizvornih govornika, izvorni su govornici neosjetljivi na morfološki oblik kolokacije, što je i očekivano s obzirom na to da je riječ o odraslim izvornim govornicima hrvatskoga kod kojih je morfološki razvoj završen.

### ***Analiza točnosti odgovora u zadatku s kolokacijama u širem kontekstu***

S ciljem ispitivanja točnosti odgovora u zadacima produktivnog znanja kolokacija u širem kontekstu provedena je dvosmjerna analiza varijance s ponovljenim mjerenjima na faktorima čestote (viša, niža) i asocijativne snage (jaka, slaba). Dobiveni su glavni efekti čestote,  $F(1,69) = 102.737$ ,  $MSE = ,036$ ,  $p < ,001$ , i asocijativne snage,  $F(1,69) = 116.197$ ,  $MSE = ,034$ ,  $p < ,001$ , te interakcija čestote i asocijativne snage na točnost odgovora,  $F(1,69) = 56.854$ ,  $MSE = ,037$ ,  $p < ,001$ . Duncanovim post hoc testom utvrđeno je kako je točnost odgovora najviša kod kolokacija više čestote i jake asocijativne snage ( $M = ,81$ ;  $SE = ,02$ ) u usporedbi s ostalim vrstama kolokacija: kolokacijama više čestote i slabe asocijativne snage ( $M = ,40$ ;  $SE = ,03$ ;  $p < ,001$ ) te niže čestote i jake ( $M = ,42$ ;  $SE = ,03$ ;  $p < ,001$ ) ili slabe asocijativne snage ( $M = ,34$ ;  $SE = ,03$ ;  $p < ,001$ ) (Slika 2), čime je potvrđena hipoteza H3.

#### Slika 2

### ***Analiza vrsta odgovora u zadacima produktivne razine***

Za odgovore u svim zadacima provedena je i analiza vrste odgovora. Odgovori ispitanika grupirani su u tri kategorije odstupanja (leksičku, gramatičku i preklapanja) unutar kojih su podijeljeni na različite vrste odstupanja. Kategorije i vrste odstupanja

prikazani su u Tablici 5. U vrstu odgovora kolokati izvan konteksta (KIK) i netočan oblik kolokata (NOK) ulazili su samo oni odgovori koji su pravopisno i semantički, a dodatno za kategoriju KIK i gramatički bili točni.

Tablica 5

Slijedi analiza kojom će se objasniti i oprmjeriti sve vrste odgovora. Analiza vrsta odgovora u zadacima produktivnoga znanja s kolokacijama u nominativu (N=35) i kosim padežima (N=35) te u širem kontekstu obuhvatila je ukupno 70 izvornih govornika hrvatskoga jezika.

### ***Analiza vrsta odgovora u zadacima s kolokacijama u nominativu i kosim padežima***

Općenito se može reći da je prosječan postotak točnih odgovora jednak neovisno o vrsti zadatka, odnosno obliku riječi (Tablica 6). Od ostalih vrsta odgovora, najzastupljeniji su kolokati izvan kontesta, pri čemu je prosječan postotak viši u zadatku s kolokacijama u nominativu (34%) u odnosu na zadatak s kolokacijama u kosim padežima (27%). U zadatku s kolokacijama u nominativu (Zadatak A) apsolutni je postotak točnih odgovora (100%) ostvaren za kolokacije *ljudska prava* i *akcijski film* dok je najniži postotak za kolokacije *statusni simbol* (11%) i *stereotipne uloge* (6%). U zadatku s kolokacijama u kosim padežima (Zadatak D) apsolutni je postotak (100%) ostvaren za kolokaciju *akcijski film* dok su kolokacije *osobni podaci* (94%) i *sportske novine* (91%) imale nešto niži postotak. Najniži je postotak točnih odgovora u zadatku s kolokacijama u kosim padežima ostvaren za kolokacije *nevladine organizacije* (26%) i *stereotipne uloge* (6%).

Tablica 6

Kako bi se dobio jasniji uvid u vrste odgovora imenskih kolokacija kod izvornih govornika hrvatskoga jezika, provedena je analiza za kolokacije različite čestote i asocijativne snage (Tablica 7).

Tablica 7

Kao što je bilo očekivano, najniži je postotak točnih odgovora za kolokacije niže čestote i slabe asocijativne snage. Najviši je postotak netočnih odgovora za kolokacije niže čestote i slabije asocijativne snage u zadatku s morfološki obilježenim kolokacijama u kosim padežima. Od ostalih vrsta odgovora, u relativno visokom postotku, osobito za kolokacije niže čestote i slabije asocijativne snage neovisno o vrsti zadatka, zastupljeni su kolokati izvan konteksta (KIK). Tako su, primjerice, ispitanici ponudili sljedeće odgovore za pojedine kolokacije:

~*ručna torba*: \*putna, \*putnička, \*manja, \*mala, \*prijenosna, \*vlastita, \*osobna, \*velika, \*sportska, \*ženska

~*stereotipne uloge*: \*poslovne, \*svakodnevnne, \*glavne, \*kućanske, \*obiteljske, \*dnevne, \*diskriminirane, \*bračne, \*društvene, \*pripisane, \*filmske, \*spolne, \*rodne



Analiza pokazuje da je postotak točnih odgovora podjednak za kolokacije više čestote i jake ili slabe asocijativne snage te niže čestote i jake asocijativne snage. Za razliku od neizvornih govornika (Ordulj, 2016) izvorni govornici imaju iznimno malen broj netočnih i gramatički nepravilnih odgovora, što je i očekivano. No, nije bilo očekivano da će i izvorni govornici, slično neizvornima, proizvesti relativno visok broj odgovora vrste “kolokacije izvan konteksta”. Navedeni rezultat pokazuje potrebu iscrpnijega istraživanja uloge konteksta u ovladavanju kolokacijama.

### **Analiza vrsta odgovora u zadatku s kolokacijama u širem kontekstu**

Prosječan postotak pojedine vrste odgovora ukazuje na podjednak postotak točnih odgovora (49%) i kolokata izvan konteksta (42%) (Tablica 8). Najviši je postotak ostvaren za kolokacije ljudska prava (97%) i životni standard (86%), a najniži je postotak primjetan za kolokacije *kupovna moć* (16%) i *smrtonosne bolesti* (7%).

#### Tablica 8

Kod dubinske analize s obzirom na čestotu i asocijativnu snagu kolokacija u širem kontekstu, uočava se visok postotak točnih odgovora kod izvornih govornika za kolokacije više čestote i jake asocijativne snage (Tablica 9).

#### Tablica 9

Za ostale je vrste kolokacija (VČ i SAS te NČ i JAS ili SAS) primjetan visok postotak vrste odgovora kolokati izvan konteksta (KIK). Pritom je za kolokacije više čestote i slabe asocijativne snage i niže čestote te jake asocijativne snage točnih odgovora podjednako kao i kolokata izvan konteksta. Neki su od primjera kolokata izvan konteksta (KIK):

~*smrtonosnih bolesti*: \*malignih, \*dobročudnih, \*neizlječivih, \*opasnih, \*autoimunih, \*drugih, \*teških, \*kroničnih, \*ostalih, \*opasnih, \*kancerogenih, \*imunih, \*čestih, \*benignih, \*prijenosnih, \*smrtnih

~*turističkim centrima*: \*trgovačkim, \*velikim, \*najboljim, \*većim, \*velikim, \*dalmatinskim, \*gradskim, \*tržnim, \*glavnim, \*urbanim, \*regionalnim

Provedena kvalitativna analiza uporabe kolokacija u širem kontekstu pokazuje da širi kontekst nema većega utjecaja na točnost odgovora, čiji je postotak usporediv s onima u užem kontekstu (izoliranim rečenicama). No, za razliku od uporabe kolokacija u izoliranim rečenicama, kod uporabe kolokacija u širem kontekstu veće su se kvalitativne razlike pokazale s obzirom na čimbenike čestote i asocijativne snage. Postotak točnih odgovora najveći je za kolokacije visoke čestote i jake asocijativne snage, a znatno niži, tek jednu trećinu, za kolokacije niske čestote i slabe asocijativne snage. Taj rezultat navodi na zaključak da bi u istraživanju ovladavanja kolokacijama valjalo ispitati odnose konteksta i različitih vrsta kolokacija.

## **Zaključak**

Dosadašnja istraživanja kolokacija u hrvatskome jeziku uglavnom su ograničena na analize kolokacija kao lingvističkih pojavnosti, no nedostaju istraživanja njihovim ovladavanjem u hrvatskome kao materinskome jeziku. Kao prvi korak u istraživanju ovladavanja kolokacijama u materinskome jeziku, istražili su se čimbenici koji utječu na uporabu kolokacija kod izvornih govornika hrvatskoga jezika. Rezultati provedenoga istraživanja navode na nekoliko bitnih zaključaka. Prvo, u uporabi kolokacija u materinskome jeziku važni su čimbenici čestota kolokacije i jačina asocijativne snage, a formalni su čimbenici (oblik kolokacije) za izvorne govornike nevažni. Razlike u točnosti uporabe kolokacija različite čestote i asocijativne snage potvrđuju postavke uporabno utemeljenih teorija ovladavanja jezikom prema kojima različite vrste čestote utječu na usvajanje jezičnih jedinica i obrazaca. Drugo, prema obilježjima uporabe kolokacija moguće je usporediti izvorne i neizvorne govornike. Izvornim su govornicima slični neizvorni govornici na visokome stupnju znanja jezika, no bitne razlike među njima ipak postoje. Kao što je već navedeno, kod izvornih je govornika nevažan čimbenik morfološkoga oblika kolokacije jer su izvorni govornici ovladali morfološkom sastavnicom jezika. Treće, utjecaj konteksta u kojem se kolokacija rabi općenito je relativno malen, no promotre li se kolokacije različitih obilježja, kontekst se pokazuje kao bitan čimbenik upravo za uporabu kolokacija niske čestote i slabe asocijativne snage, što je iznimno važan nalaz upravo za poučavanje jezika. S obzirom na to da je istraživanje pokazalo relativno velik postotak proizvedenih kolokacija koje ne odgovaraju zadanom kontekstu (premda su ovjerene kolokacije), važno bi bilo istražiti (i opisati) odnos kolokacije i konteksta u kojem su pojavljuje. Rezultati provedenoga istraživanja važni su ne samo zato što daju uvid u opće procese ovladavanja hrvatskim kao materinskim jezikom i njegove uporabe, već mogu biti polazište za promišljanje o novim načinima i učinkovitijim postupcima njegova poučavanja u obrazovnome sustavu.