

Distribution of verbal overgeneralizations in the Serbian Corpus of Early Child Language¹

Darinka Anđelković²

*Laboratory of Experimental Psychology and Institute of Psychology
Faculty of Philosophy, University of Belgrade*

Mirjana Mirić

Institute for Balkan Studies, Serbian Academy of Sciences and Arts

*The study is aimed at exploring the occurrence of innovative verb forms recorded in early spontaneous children's production of Serbian – a language with rich inflectional and derivational morphology. The overgeneralized verbs were retrieved from the corpus of eight children's production, longitudinally recorded from 1;6 to 4;0, and the developmental patterns of the distribution across age were explored. The analysis shows that overgeneralizations *hoćem 'want' 1.sg.pres., *nećem 'want' 1.sg.pres., *možem 'can' 1.sg.pres. and *bidem 'be' 1.sg.pres. are the typical representatives of early overgeneralizations in Serbian. They are typically recorded in all children's speech samples with relatively high frequency. The overgeneralized *hoćem and *nećem developmentally precede all other overgeneralized verbal forms, but disappear earlier than others. The overgeneralizations of contentive verbs, usually hapaxes, spread later and retain their position long after. The findings are discussed in comparison with the previous findings on overgeneralizations, with particular attention to the developmental patterns they exhibit.*

Keywords: overgeneralization, word formation, child language corpus, verbal morphology, the Serbian language

The aim of the present study is to describe the occurrence and developmental patterns in the production of verbal morphological overgeneralizations in the course of acquisition of a morphologically rich

1 The study was supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia, grant numbers ON179033 and 178010.

2 Corresponding author: dandjelk@f.bg.ac.rs

language. In linguistics, overgeneralizations are typically defined as the application of a grammatical rule in cases where it does not apply (Nordquist, 2017). In contrast to analytical languages, such as English, Serbian is a fusional language with a highly developed inflectional and derivational morphology. Since both are the most important sources of structural information and meaning, it is worthwhile to explore a sample of children's innovative word forms at the moment of their composition. Overgeneralizations in early language production reveal that a child recognizes a meaning as appertaining to a certain class of words, wields some knowledge on structural regularities in his/her first language, and uses the regularities in his/her efforts of a new-word formation. Overgeneralizations in spontaneous children's production uncover the connection between the form and function at the moment of its emergence, which reveals the regularities known to a child, and aspects of mother tongue grammar that are still to be acquired.

Overgeneralizations in early child language have attracted a lot of attention in the field of language acquisition (Albright & Hayes, 2003; Ambridge, Freudenthal, Pine, Mills, Clark, & Rowland, 2009; Ambridge, 2010; Ambridge, Pine, Rowland, Chang, & Bidgood, 2013; Anđel, Klampfer, Kilani-Schoch, Dressler, & Kovačević, 2000; Bowerman, 1988; Brooks, Tomasello, Dodson, & Lewis, 1999; Brown, 1973; Katičić, 2003; Kuczaj, 1977; Hržica, 2012; Li & MacWhinney, 1996; MacWhinney, 1976; MacWhinney, 1993; Marcus, Pinker, Ullman, Hollander, Rosen, & Xu, 1992; Slobin, 1973; Perek & Goldberg, 2015; Theakston, 2004). In the domain of verbal morphological overgeneralizations, the typically discussed cases are overregularization errors of the English Past Tense *-ed* and the German regular *-t* participle, with the overregularization rate of 4% to 5% (Clahsen & Rothweiler, 1993; Markus et al., 1992). The early observed and thoroughly described English Past Tense suffix *-ed* is usually overgeneralized in early child language in two different ways. One is adding *-ed* to the generic form of an irregular verb (*eated*, *goed*), while another is attaching the same suffix to the past tense form of an irregular verb (*ated*, *wented*) (Brown, 1973; Ervin, 1964; Kuczaj, 1977; Slobin, 1973). It was also reported that the former had developmental priority, while the latter appeared after the irregular *went* and *ate* had already been acquired. Both the correct and incorrect forms were often used simultaneously in development, which indicates that a child must handle opposing information from the input and learn that the irregular past tense forms deviate from the general rule without much explicit aid from adults.

Considered related to the question of negative evidence and the so-called "logical problem of language acquisition" (LPLA) (Baker 1979; Baker & MacCarthy, 1981; Pinker, 1984), overgeneralizations have become a battlefield for significant theoretical dispute in the recent decades. In generative linguistics, children's productivity is considered to be the result of abstract knowledge that moulds the new coming words into the rules of the input language. In efforts

to explain the retreat of overgeneralizations in the course of development, the thesis of direct blocking of overgeneralization by the corresponding irregular form was offered and supported in the field (Anderson, 1977; Baker, 1979; Baker & MacCarthy, 1981; Clark, 1983; Pinker, 1984). However, it was also opposed by a more general Competition Model (MacWhinney, 1988; 1989; 1993), in which overgeneralizations are considered as subject to three types of pressures: the analogic pressure that rouses overgeneralizations, the auditory representations from the episodic support that reigns them back, and the competition between them. The role of analogy was also suggested in the single-route model (Bybee & Moder, 1983) and supported in the connectionist accounts (Albright & Hayes, 2003; Ambridge, 2010; Rumelhart & MacClelland, 1986; Skousen, 2001), which advocated that rules are superfluous if parallel distributed processing is proposed.

The debate referred primarily to the English grammatical forms, so the models offered for one language might not be necessarily applicable to other languages. Important insights into the verbal overgeneralizations in a morphologically complex language were provided by a study of early Spanish verbal overgeneralization errors, which has shown that overgeneralizations are possible in both the stem/root and suffixes (Clahsen, Avelo, & Roca, 2002). Children overapplied regular forms to irregular, but not vice versa, with the onset linked to the appearance of obligatory finiteness markings. The low frequency irregular verb forms yielded more errors than the high-frequency ones. Similarly, Aguirre (2003) reported that one particular class of verbs in Spanish (the most salient one) became preferable once the studied child entered the proto-morphology stage.

The significance of language typology was emphasized in the framework of Natural Morphology (Dressler, 1985; Dressler, 2005), proposing that grammatical modules are not innate, and that children gradually learn the grammar from the input. The children acquiring a system with rich morphology are more tuned to morphology than the children acquiring the languages of simple morphology. The typological properties of a particular language are essential for the course of development because they are acquired already at the proto-morphological phase. Productivity and transparent morphology facilitate the early start of the usage of language specific structural properties (Bittner, Dressler, & Kilani-Schoch, 2003; Dressler, 1985; Dressler, 2005; Hržica, 2011; 2012; Katičić, 2003; Radisavljević, 2013).

Several studies on Croatian are of particular importance for our research, given the typological similarity between Serbian and Croatian (Anđel et al., 2000; Hržica, 2012; Katičić, 2003). The main findings of studies on Croatian are the analogy error found in the 1.sg.pres.³ of the verb *htjeti* 'want'

3 Abbreviations used in the paper: sg - singular, pl - plural, f - feminine, m - masculine, inf - infinitive, pres - present tense, fut - future tense, imp - imperative, ppart - past participle,

(**hoćem* < *hoću*) across all studies, and the phenomenon of the ‘class shift’ (Katičić, 2003) or ‘overregularization’ (Hržica, 2012), in which the transparent and productive forms substituted the opaque ones (e.g. **pisam* < *pišem*).

The studies in morphologically rich languages cited above varied from spontaneous production to narratives, and were conducted on a very small number of children (1 to 3), except for Clahsen et al. (2002). In addition, it is quite difficult to compare the distribution of the first occurrence and developmental patterns of overgeneralizations across age levels, given that the studies differed regarding the age span, e.g. 1;7–1;10 (Aguirre, 2003); 1;10–2;2 (Anđel et al., 2000); 1;2–3;2 (Hržica, 2012). The phenomenon of overgeneralizations in early child language is yet to be explored in Serbian, which makes the description of their first occurrence and distribution in a larger number of children highly significant. The Serbian Corpus of Early Child Language enables a longitudinal exploration of the spontaneous production of a relatively large number of children (8) in a relatively wide age span (16 age levels between 1;6–4;0).

Verbal morphology in Serbian

Before we present our study, a simplified overview of Serbian verbal morphology is necessary. The verbs in Serbian are generally inflected for person (1st, 2nd, 3rd), number, and gender (marked in participles only), as well as tense and mood.⁴ They consist of a stem and an inflection, where the stem comprises a root and a thematic vowel, as in (1):

(1) <i>pevam</i>	[<i>pev-</i> + <i>-a-</i>]	+	<i>-m</i>
‘sing’	[ROOT + THEMATIC VOWEL]stem	+	INFLECTION
1.sg.pres.			1.sg.pres.

Thematic vowels and consonants intervene between the root and the inflection (e.g. *pev-a-ti* ‘to sing’, *misl-i-ti* ‘to think’, *trep-nu-ti* ‘to blink’). They are responsible for the choice of a conjugation pattern a given verb belongs to (Radisavljević, 2013).

In order to make a verbal form in Serbian, the inflectional ending is added onto the present or infinitival stem. The present stem is formed by omitting the inflectional ending *-mo* from the 1.pl.pres. (e.g. *pev-a-mo* ‘sing’ 1.pl.pres.

dim - diminutive, refl - reflexive, neg - negative, imperf - imperfective aspect, perf - perfective aspect.

4 Finite tenses are present, aorist, imperfect (synthetic), perfect, pluperfect, future I (analytic); mood forms are imperative, conditional and future II; infinite forms are infinitive, active and passive past participle, present and past verbal adverbs. Among tenses, only present, perfect and future I are productive in Serbian (Radisavljević, 2013: 51). This fact has been confirmed in the language sample of four children in the SCECL (age 1;6-3;0) (Mandić, 2013: 154).

> *peva-*). There are two ways of making the infinitival stem – by omitting the infinitival marker *-ti* in the verbs ending in *-ti* preceded by a vowel (e.g. *pev-a-ti* ‘sing’ inf. > *peva-*) or by omitting the inflectional ending *-oh* from the 1.sg.aorist in the verbs whose infinitive ends in *-ti* preceded by a consonant (*tresti* ‘shake’ inf. > *tres-oh* 1.sg.aorist > *tres-*) or in *-ći* (*seći* ‘cut’ > *sek-oh* 1.sg.aorist > *sek-*) (Stanojčić & Popović, 2000: 108). Among the verbal forms relevant for this paper, the present and imperative are built from the present stem, whereas the past participle (used in the perfect) and synthetic future are built from the infinitival stem. The examples of forms which emerge in children’s production are given below, exemplified by the verb *pevati* ‘sing’.

Present tense			Future tense (synthetic)		
	sg.	pl.		sg.	pl.
1.	pev-a-m ⁵	pev-a-mo		pev-a-ću	pev-a-ćemo
2.	pev-a-š	pev-a-te		pev-a-ćeš	pev-a-ćete
3.	pev-a	pev-a-ju		pev-a-će	pev-a-će
Imperative			Past participle		
	sg.	pl.		sg.	pl.
1.	/	pev-a-j-mo	m.	pev-a-o	pev-a-li
2.	pev-a-j	pev-a-j-te	f.	pev-a-la	pev-a-le
3.	/	/	n.	pev-a-lo	pev-a-la

According to the traditional Serbian grammars, there are seven conjugational classes, differing in morphological complexity (Table 1).

Table 1. Conjugational classes in Serbian (adapted according to Stevanović, 1986; Stanojčić & Popović, 2000: 109–113)

Conjugational class	Present stem ending	Infinitival stem ending	Example
I	-e (trese-)	consonant (tres-)	<i>tresti</i> ‘shake’
II	-e (piše-) ⁶	-a (pisa-)	<i>pisati</i> ‘write’
III	-ne (brine-)	-nu (brinu-)	<i>brinuti</i> ‘worry’
IV _A	-je (čuje-)	-Ø (ču-)	<i>čuti</i> ‘hear’
IV _B	-je (kuje-)	-a (kova-)	<i>kovati</i> ‘mint / coin’
V _A	-ā ² (pevā-)	-a (peva-)	<i>pevati</i> ‘sing’
V _B	-ē (razumē-)	-e (razume-)	<i>razumeti</i> ‘understand’
VI	-ī (nosi-, vidi-)	-i (nosi-), -e (vide-)	<i>nositi</i> ‘carry’, <i>videti</i> ‘see’
VII	-i (drži-)	-a (drža-)	<i>držati</i> ‘hold’

5 The inflection *-u* in 1.sg.pres. is used only with three verbs: *hoć-u* (1.sg.pres. ‘I want’), *neć-u* (1.sg.pres. ‘I do not want’), *mog-u* (1.sg.pres. ‘I can’). All other verbs bear the ending *-m* in 1.sg.pres.

6 Resulting in a consonant mutation.

7 Long vowels -ā, -ē and ī.

A different approach towards the inflectional verbal classes in Serbian is taken by Radisavljević (2013). Following the model of Natural Morphology proposed in the classification of verbs in Croatian (Dressler, Dziubalska-Kořaczuk, & Katićić 1996), the author describes the most productive verbal microclasses in Serbian in terms of inflectional productivity.

Apart from the rich inflectional morphology, Serbian has a productive system of derivational morphology, which in case of verbs serves as an aspectual marker. Traditionally, Serbian verbs are divided in two classes: perfective and imperfective. According to Arsenijević (2006: 202): “the stem verb is normally imperfective [...]. Adding a prefix to a stem verb contributes a lexical meaning (often even causing a shift in the lexical meaning of a verb), and it makes the verb perfective. [...] Adding a suffix to a perfective verb (even to a perfective stem verb) makes the verb imperfective. The suffix does not contribute any lexical meaning”. For example, *pevati*_{imperf} ‘sing’, *zapevati*_{perf} ‘start singing’, *otpevati*_{perf} ‘finish singing’.

Aims

Having in mind the findings of previous studies in different languages (e.g. Albright & Hayes, 2003; Anđel et al., 2000; Bittner et al., 2003; Bybee & Moder, 1983; Hržica, 2012; MacWhinney, 1988; 1989; 1993; Rumelhart & MacClelland, 1986; Skousen, 2001), the following distributional regularities in Serbian overgeneralizations are to be expected:

- a. First verbal overgeneralizations are highly expected at the earliest age level of the Serbian corpus (1;6). Since they announce the beginning of grammatical development, the onset of overgeneralizations may appear at different age levels in individual children.
- b. The most prevalent form of overgeneralizations is the overregularization of irregular forms (cf. Clahsen et al., 2002; Hržica, 2012).
- c. The role of analogy in the overgeneralizations (e.g. MacWhinney, 1988; 1989; 1993) is expected, in terms of using inflections or thematic vowels analogous to other forms or verb classes evident from the input (cf. Hržica, 2012; Katićić, 2003).
- d. Overgeneralizations are a temporary phenomenon in the acquisition of verbal morphology, and it is possible to track the onset, prevalence and ending point of their occurrence in the children’s production.

The appearance of overgeneralizations and neologisms in early production of a synthetic language reveals the developmental changes in the perception of words from the language input, and uncovers the developmental beginning of word parsing. An exploration of the formal characteristics provides insight into the parsing of words (root vs. derivation vs. inflection), which are generally perceived as units. Due to the scope of the paper, which is restricted

to the distributional properties of early overgeneralizations in Serbian, formal characteristics are going to be only briefly discussed, while a detailed linguistic analysis will be provided in a prospective article (Authors, in preparation).

Method

The data for the study were taken from the *Serbian Corpus of Early Child Language (SCECL)* (Anđelković, Ševa, & Moskovljević, 2001) in *The CHILDES Database* (MacWhinney & Snow, 1985).

The corpus consists of app. 1.000.000 words in total, out of which more than 235.000 have been produced by children. An automatic lemmatization of words was conducted on the basis of the *Frequency Dictionary of Contemporary Serbian Language* (Kostić, 1999), which is a procedure that produced only 4% of errors in the case of adult language (Ilić & Kostić, 2003a, 2003b). When a sample of children's production is targeted, the phonological deviations from the conventional form of words request for additional manual work and annotation, especially in the samples of young ages. The CLAN programs *FREQ* and *KWAL* of the CHILDES Project were used in order to retrieve the children's verbs and check for their linguistic and situational context. The analysis of the verbal forms properties revealed overgeneralizations as innovative categories, which will be concisely described. The distribution of their usage was inspected across age levels and discussed from the developmental point of view.

The sample for the corpus consisted of eight monolingual Serbian children (4 boys, and 4 girls), whose spontaneous speech production was longitudinally recorded from 18 to 48 months of age. The recordings were made at 16 age levels, approximately at 2-month intervals, and lasted for 90 minutes each. The children belonged to the families of middle SES, living in Belgrade and Banja Luka, equally distributed. Three children had an older sibling, five were single. In four families, at least one parent was highly educated (university or college), while all others had a high school degree.

Table 2 presents the basic measures of language development in the sample: a. the Mean Length of Utterances calculated by words MLUw (in order to exclude numerous exclamations, non-linguistic expressions, and simple nominations, only utterances containing verbs were included); b. vocabulary size estimated by the average number of types (lexical entries) at all ages; c. vocabulary size (N of types) at the earliest age level (1;6).

Table 2. Indications of language development in individual children

	ANA	ANE	DAC	JEL	LAZ	LUK	MIL	NIK
MLUw	3.289	2.722	2.899	2.514	2.509	3.144	2.035	3.157
Average N of types 1,6 - 4;0	323	267.62	229.12	209.87	224.62	341.12	151.25	265.31
N of types at the age of 1;6	156	68	89	56	73	193	25	61

An example of an overgeneralized form used in the context is given below.

(2) *nećem* (JEL2;0)

- *MAJ: pokaži Darku gde je nos .
 %mor: v|pokazati n:prop|Darku conj|gde v|biti n|nos .
 %eng: 'Show Darko where the nose is.'
 *JEL: 0 .
 %act: pokazuje nos
 %eng: showing the nose
 *MAJ: a gde je oko ?
 %mor: conj|a conj|gde v|biti n|oko ?
 %eng: and where is the eye
 *JEL: **nećem@z:nv !**
 %mor: v:neo|nećem !
 %eng: 'I do not want to.'
 %act: uzima slikovnicu sa stola
 %eng: taking the picture book from the table

Results

1. Prevalence of verbal overgeneralizations in the SCECL

An approximate prevalence of children's overgeneralizations estimated on the automatically lemmatized corpus is presented in Table 3. A manual itemized analysis has been performed on verbs, while other parts of speech are to be explored in detail in the future. The main reason is the fact that verb overgeneralizations are an extensive and quite productive group of overgeneralization errors, which effectively reflect the complexity of the grammar.

Table 3. An approximate prevalence of children's overgeneralizations in the SCECL (age span 1;6–4;0)

	N	% of the overall number of types and tokens ⁸
Overgeneralizations (types)	140*	2,7*
Overgeneralizations (tokens)	300*	0,15*
Verbal overgeneralizations (types)	61	4,9*
Verbal overgeneralizations (tokens)	126	0,21*

*The approximate values

8 Types – different verb forms per lemma, token – occurrence for each specific type.

Overgeneralizations retrieved from the corpus belong to the following categories of verbs: modals (e.g. *moći* 'can'), auxiliaries (e.g. *biti* 'be') and contentive (lexical) verbs (e.g. *doneti* 'bring').

The exact number of reported overgeneralizations among verbs (calculated in types) in the child language of the SCECL is 61, with the overall frequency of 126. Having in mind the overall children's production, it is obviously a rare phenomenon in Serbian, just as it is in other languages. Nevertheless, the information that every twentieth verb calculated in the types (4.9%) in early child language is overgeneralized illustrates the importance of the analysis that follows.

2. Morphological properties of early verbal overgeneralizations

The inspection of inflectional and derivational properties of overgeneralized verbs revealed four relatively homogenous categories of overgeneralization errors recorded in spontaneous Serbian children's production: stem overgeneralization, inflectional overgeneralization, changes in derivation and compounding, and idiosyncratic verbs (neologisms). Having in mind the lack of studies of overgeneralizations that would report on the difference between various types of overgeneralization errors, the categories reported herein were mainly differentiated by relying on the method of induction. The primary criterion was to explore whether the errors were made at different formative parts of the verb: the inflection, the stem, the prefix or the suffix.

a. **Stem overgeneralizations.** This category contains two subcategories. Firstly, the stem of another verbal form of the same verb is overgeneralized, while the inflectional ending of the new form remains typical for the given form. The most prominent examples in this group are **bidem* < *budem* ('be' 1.sg.pres.), **bideš* < *budeš* ('be' 2.sg.pres.), and **bide* < *bude* ('be' 3.sg.pres.). The root of the stem was produced by analogy with the infinitival stem *bi-* in the infinitive *bi-ti* and/or past participle *bi-o*, *bi-la*, *bi-lo*, etc.⁹

Apart from the auxiliary *biti* 'be', numerous contentive verbs also exhibit this type of stem overgeneralization: **donesti* < *doneti* ('bring' inf.), **odnesećemo* < *odnećemo* ('take away' 1.pl.fut.), **popnićeš* < *popćeš* ('climb' 2.sg.fut.), **kaziła* < *kazala* ('tell' f.sg.ppart.), etc., which in the last case leads to the consonant mutation. The typical pattern observed within this category is the use of the present stem instead of the infinitival one of the same verb, and vice versa.

The second type of stem overgeneralization has been registered in the verbs in which a thematic vowel is substituted by a vowel typical for some other conjugational class: **donesam* (analogous to class V_A*pevam*) < *donesem*

9 The verb *biti* 'be' shows a high degree of irregularity throughout the whole paradigm.

(‘bring’ 1.sg.pres.), **izmišljem* (analogous to class I *tresem*) < *izmišlj \underline{m}* (‘think up’ 1.sg.pres.), **namazam* (analogous to V_{Λ} *pevam*) < *namažem* (‘spread on’ 1.sg.pres.). It is this type of overgeneralizations which is typically studied in morphologically complex languages (cf. Anđel et al., 2000; Hržica, 2012).

b. **Inflectional overgeneralizations.** This category includes the cases of an inappropriate inflection added to a verb root by analogy with other inflected verb forms.¹⁰ Typical examples are the forms **hoć-e- \underline{m}* , **hoć-a- \underline{m}* , or **hoć-i- \underline{m}* ‘want’ 1.sg.pres. used instead of the target form *hoć- \underline{u}* .¹¹ It is also the case with **neć-e- \underline{m}* ‘want.neg’ 1.sg.pres. < *neć- \underline{u}* , and **mož-e- \underline{m}* or **mog-a- \underline{m}* ‘can’ 1.sg.pres. < *mog- \underline{u}* . These are modal/auxiliary verbs which deviate from the Serbian verb paradigm in 1.sg.pres., given that these three verbs are the only ones to bear the ending *-u* instead of *-m* in that form. This irregularity probably causes the children’s inclination to use the more frequent inflection *-m* instead of the low frequent *-u*.

The analysis has also revealed certain patterns of analogy: e.g. with the conjugational class I (*tresem*) in the case of **hoćem* and **možem*, the conjugational class VI (*nosim*) in the case of **hoćim*, or the class V_{Λ} (*pevam*) in the case of **hoćam* and **mogam*. Therefore, these examples are not just the cases of an inflectional, but also the stem overgeneralization. As will be shown later, this small group of verbs makes the most frequent kind of overgeneralization in Serbian children’s production.

Errors have also been found in the case of 3.pl.pres. and 2.sg. imp. of the contentive verbs, e.g. **ljub-a- \underline{ju}* < *ljub- \underline{e}* (‘kiss’ 3.pl.pres.), **ručk- \underline{e}* < *ručk-a- \underline{ju}* (‘eat lunch’ 3.pl.pres.dim.),¹² **doručk-a- \underline{j}* instead of *doručk- \underline{u} - \underline{j}* (‘eat breakfast’ 2.sg.imp.).

c. **Changes in derivation and compounding.** This category of overgeneralization relies on highly productive derivational tools of prefixation and suffixation for the new verb formation. The root of the verb remains unchanged, as well as the inflectional morphemes. Several subcategories of

10 The cases in which the form of the 3.sg.pres. is used instead of other person and number forms of present tense has not been taken into account in this paper, given that it presents the inflection omission. The form of 3.sg.pres. lacks an inflection - it is marked only with the stem vowel. This type of error in early child Serbian was analyzed in Mandić (2013), on the sample of four children from the SCECL as a default verb form in the earliest utterances – analogous to the phenomenon of root infinitives in other languages (cf. Katičić, 2003 for a similar type of analysis in Croatian and Aguirre, 2003 for Spanish).

11 Only the full forms of the verb *hteti* ‘want’ were overgeneralized, while the enclitic forms of this verb (typically used to build the analytic future I tense) are not considered. The verb *hteti* ‘want’ shows a high degree of irregularity throughout the whole paradigm.

12 In the case of the overgeneralization **ručke* and the target form *ručkaju*, it is noteworthy that both forms could be considered as diminutive forms, which are typical in the child directed speech and child language, and in some cases become family words.

derivational changes have been recorded: prefix omission – *ključaj* instead of *zaključaj/otključaj* ('lock/unlock' 2.sg.imp.), prefix addition – *posređujemo* < *sređujemo* ('put in order' 1.pl.pres.), prefix substitution – *popolovili* < *prepolovili* ('halve' m.pl.ppart.), suffix substitution – *crtavati* ('draw' inf.) < *crtati*, etc.

d. **Idiosyncratic verbs (neologisms)**. Some children are also inclined to create the verbs on their own by inventing them from scratch. The meaning is often not clear since the verb root is idiosyncratic, but they are intriguing because of the regularity of the morphological form which completely reflects the Serbian conjugational system (e.g. *tać-i-m* 1.sg.pres., *duš-a-š* 2.sg.pres., *dibav-i-m* 1.sg.pres.), and in some cases productive derivational morphology (e.g. *is-viliću*, *za-pući*). A few overgeneralized verbs have a semi-transparent meaning: *zamejuče se* ('turn on to start playing' 3.sg.pres.refl.), *ovršili* ('to touch /the throat/ with the top of a stick' m.pl.ppart.), *zabalaiš* ('stuck' 2.sg.pres.).

A more detailed analysis of the formal properties of all categories of overgeneralizations described above is to be provided in a future study (Authors, in preparation).

Developmental patterns

Although the overall number and frequency of overgeneralizations does not allow us to draw conclusions on the underlying language acquisition mechanisms, the data provide an opportunity to search for the patterns in usage and point out the relevant developmental regularities.

All eight children in the corpus used overgeneralized forms; however, some children appear to be more productive when it comes to the number and frequency of various overgeneralized forms and the number of different categories (cf. Table 4).

Table 4. Number and frequency of overgeneralizations

Child	Number of types	Overall frequency	Overgeneralization categories in the usage*
DAC	3	6	b, c, d
NIK	5	8	a, b, d
LAZ	6	14	a, c, d
MIL	8	13	a, b
ANE	11	13	a, b, c, d
LUK	12	17	a, b, c
JEL	14	28	a, b, c
ANA	18	28	a, b, c, d

* a – stem overgeneralization, b – inflectional overgeneralization, c – changes in derivation and compounding, d – idiosyncratic verbs (neologisms)

Even though the earliest instances of overgeneralizations are found already at the youngest age in the sample (1;6), it is only a property of the two most talkative children: ANA and LUK have the largest vocabulary size in the sample and are among the three children that produce the longest utterances (Table 2). The others start with overgeneralizations later (JEL 1;8, ANE 1;10, LAZ 2;2, NIK 2;2, DAC 2;10, MIL 2;10), which shows that they do not appear at the stage of children's earliest utterances, but probably only with the onset of grammatical marking, which will be explored in more detail in the future.

Remarkably, the earliest age levels of all children (until 2;0) lack variety in the overgeneralized forms, given that only two verbs are overgeneralized: **hoćem* 'want' 1.sg.pres. and **nećem* 'want' 1.sg.pres.neg., always in the same grammatical form.¹³ Even if children start using overgeneralizations at later ages, their first errors exhibit the same pattern typically affecting the modal *hteti* 'want' and *ne hteti* 'want.neg.' (**hoćem* JEL 1;8, ANE 2;0, DAC 2;10, MIL 2;10; **nećem* ANA 1;6, JEL 1;8, ANE 1;10, NIK 2;8). Moreover, these verbs are overgeneralized in the speech samples of most children (*hteti* 'want' in 6, *ne hteti* 'want.neg.' in 4 children). Along with the overgeneralized form of the modal *moći* 'can' – **možem* recorded in 6 children, and the forms of auxiliary *biti* 'be' – **bidem*, **bideš*, **bide* produced by 4 children, they are the most frequently overgeneralized verbs in the sample (*hteti* 23, *ne hteti* 7, *moći* 10, *biti* 20). Therefore, this kind of inflectional overgeneralizations, and, moreover, the tendency to regularize the irregular verb forms, might be taken as a typical form of verbal overgeneralizations in early Serbian child language.

All other overgeneralizations, affecting the contentive verbs and creation of neologisms, are recorded in very low frequency or as hapaxes. They make an open group of verbs from different conjugational classes traced in children, which shows that the search for proper conjugation and morphological regularities is applied. Their occurrence stretches towards the latest age levels, probably for as long as the vocabulary grows.

Furthermore, the data reveal that overgeneralizations of modal/auxiliary verbs *hteti* 'want', *ne hteti* 'want.neg.', *moći* 'can' and *biti* 'be' typically precede in development the overgeneralizations of contentive verbs and neologisms (Table 5).

13 The exceptions are two verbs also reported at an early age: **donesti* 'bring' inf. (ANA 2;0) and *ključaj* '(un)lock' 2.sg.imp. (JEL 1;10), and their token frequency is 1.

Table 5. The age of the first appearance of the overgeneralized forms of modal/auxiliary verbs, contentive verbs and neologisms

	Modal/auxiliary verbs	Age	Contentive verbs	Age	Neologisms	Age
ANA	<i>hoćem</i> 'want' 1sg.pres.	1;6	<i>opirne</i> 'blow' 3.sg.pres.	2;0	<i>dušaš</i> 'unknown' 2.sg.pres.	2;8
ANE	<i>nećem</i> 'want' 1.sg.pres.neg.	1;10	<i>obuciti</i> 'dress' inf.	2;2	<i>dibavam</i> 'unknown' 1.sg.pres.	2;2
DAC	<i>hoćem</i> 'want' 1sg.pres.	2;10	<i>isipam</i> 'pour' 1.sg.pres.	3;2	<i>zabalaiš</i> 'stuck' 2.sg.pres.	2;8
JEL	<i>hoćem</i> 'want' 1sg.pres.	1;8	<i>ključaj</i> 'lock/unlock' 2.sg.imp.	2;0	/	/
LAZ	<i>bide</i> 'be' 3.sg.pres.	2;2	<i>popolovili</i> 'halve' m.pl.ppart.	3;0	<i>taćim</i> 'unknown' 1.sg.pres.	2;2
LUK	<i>hoćem</i> 'want' 1.sg.pres.	1;6	<i>zapravim</i> 'repair' 1.sg.pres.	2;8	/	/
MIL	<i>hoćem</i> 'want' 1.sg.pres.	2;10	<i>uzmo</i> 'take' m.sg.ppart.	3;2	/	/
NIK	<i>nećem</i> 'want' 1.sg.pres.neg.	2;8	<i>ljubaju</i> 'kiss' 3.sg.pres	2;2	<i>isviliću</i> 'unknown' 1.sg.fut.	3;8

The only exception is NIK, whose samples provide the first instance of the contentive verb **ljubaju* 'kiss' 3.pl.pres. chronologically before the first instance of the overgeneralized modal **nećem* 'want' 1.sg.pres.neg. at the age of 2;8. However, it is possible that the overgeneralized forms of modal/auxiliary verbs have simply not been recorded in his production at early ages due to the limited duration of the recording sessions.

In addition, it is noteworthy that the overgeneralized **hoćem* and **nećem* are not reported in the sample after 3;2 in any child, which suggests that after the age of 3 their conjugation has become stable and target-like. On the other hand, the overgeneralized forms of *moći* 'can' and *biti* 'be', and various contentive verbs and neologisms have been recorded even at the latest age levels in the sample (4;0) (Table 6).

Table 6. The age span of the usage of the overgeneralized forms of verbs

	Modal <i>hteti</i> 'want' <i>ne hteti</i> 'want.neg.'	Modal/auxiliary <i>moći</i> 'can' <i>biti</i> 'be'	Contentive verbs	Neologisms
ANA	1;6–2;0	2;6	2;0–2;10	2;8–3;8
ANE	1;10–2;0	2;4	2;2–4;0	2;2–3;2
DAC	2;10–3;2	3;0–3;10	2;4	2;8
JEL	1;8–2;4	2;4–4;0	2;0–3;0	/
LAZ	/	2;2–3;4	3;0	2;2–2;4
LUK	1;6–2;10	2;6–3;0	2;8–4;0	/
MIL	2;10	3;8	3;2–4;0	/
NIK	2;8	3;0	/	3;8

Discussion

The results presented above seem to support the findings from the previous studies and improve our knowledge on overgeneralizations in fusional languages.

The overall rate of verbal overgeneralizations in early Serbian (4.9%) is in accordance with the rates obtained for English and German (cf. Markus et al., 1992 for English and Clahsen & Rothweiler, 1993 for German), and slightly higher than the rate of the overregularization errors in Spanish (2%), although we must bear in mind that these figures refer to the instances of overregularization as one type of overgeneralization errors.

With regard to the categories of overgeneralization errors, the following tendencies are observed.

In the domain of inflectional morphology, children tend to regularize irregular forms. There are verbal forms which seem particularly difficult for children, such as *hoću* 'want' 1.sg.pres., *neću* 'want' 1.sg.pres.neg. and *moгу* 'can' 1.sg.pres., given that the vast majority of errors involves these verbs. The most plausible explanation for this finding is the fact that these three verbs bear a different inflection (*-u*) in the 1.sg.pres. in comparison to all other verbs in Serbian which bear *-m*. Therefore, the errors represent a tendency to conjugate the verbs by analogy with the other inflected forms. Exactly the same error (**hoćem < hoću*) was reported in Croatian (Hržica, 2012; Katičić, 2003). The issues of both overregularization and frequency have already been pointed out in previous research (Ambridge, 2010; Brown, 1973; Clahsen et al., 2002; Hržica, 2012; Katičić, 2003; Kuczaj, 1977; MacWhinney, 1976; MacWhinney, 1993; Marcus et al., 1992; Slobin, 1971, 1973; Perek & Goldberg, 2015). Regarding the frequency, it is noteworthy that *hteti*, *ne hteti* and *moći* are high-token-frequency verbs in Serbian. However, the inflectional ending *-u* in 1.sg.pres. has been found in these three verbs only, which makes the form ending in *-u* infrequent and non-productive.

The categories of overgeneralizations we reported herein match the categories of overgeneralizations recorded in early Croatian, at least to a certain extent – in the domain of analogy errors reported in the use of the forms **hoćem < hoću* (Anđel et al., 2000; Hržica, 2012; Katičić, 2000), and the stem overgeneralization – the case when a child opts for a wrong stem, e.g. **pis-a-m < piš-e-m* (Hržica, 2012). These findings suggest that typologically similar languages exhibit similar patterns of overgeneralizations. However, unlike the studies on Croatian which mainly focused on one type of errors (i.e. overregularization or class shift), the data from the SCECL revealed two other types of overgeneralization errors: the changes in derivation and compounding and neologisms. All types of overgeneralizations deserve a more detailed analysis in the future, especially in relation to the frequency of

the overgeneralized inflections, thematic vowels, prefixes and suffixes on the verb forms in the input. An additional analysis of stem overgeneralizations (and possibly a re-examination of the verbal classes in Serbian) should be conducted in order to provide the data comparable to the studies conducted on other morphologically rich languages and to test the model of Natural Morphology (Bittner et al., 2003) on Serbian.

It is important to emphasize that our findings differentiate between *hteti* 'want' and *ne hteti* 'want.neg.' on one side, and all other overgeneralizations on the other. It seems that *hteti* 'want' and *ne hteti* 'want.neg.' are the first target of children's generalization of the grammar perceived in the input, and the first carriers of morphological regularities among Serbian verbs. *Moći* 'can' and *biti* 'be', even though they belong to the same group of highly frequent modal/auxiliaries, are only later affected by overgeneralization, and therefore do not have such a prominent function in the course of language acquisition. On the other hand, from the developmental point of view, they seem to resemble contentive verbs that appear a little bit later in a large number and low frequency and are affected by overgeneralization long after. This layout reveals the linkage between the gradual growth of vocabulary and the acquisition of conjugational rules.

Regarding the age span, the study showed that certain types of overgeneralizations continue to exist even at later ages (at least until the age of 4;0, as covered by the SCECL). Therefore, it is of high importance to include a wide age span when studying this phenomenon.

The analysis of morphological properties of overgeneralizations potentiated the visibility of the following processes in new word formation:

- a) Classification. Children recognize quite a complex range of Serbian verbal forms and conjugational classes early. This supports the findings on an early sensibility to specific structural properties of morphologically rich languages (Dressler, 1985; 2005; Hržica, 2011).
- b) Regularization. The recognized patterns are used in the production of grammatical forms which causes a regularization of irregular forms.
- c) Analogy. The main source of overgeneralization is analogy. This supports the previous findings on its role (Albright & Hayes, 2003; Bybee & Moder, 1983; Katičić, 2003; MacWhinney, 1988; 1989; 1993; Rumelhart & MacClelland, 1986; Skousen, 2001). The newly coined verbs are inevitably guided by the language specific patterns of verb formation which pressure overgeneralizations to arise, and all parts of a verb are potentially affected – the inflectional morpheme, the stem or the thematic vowel.
- d) Semantic differentiation. If semantic differentiation is needed in a particular context, Serbian children use all morphological tools

available in the input – prefixation, suffixation, stem alternation, which provide the basis for the construction of overgeneralizations of any part of a word.

- e) Lexicalization of unconventional meanings. In communication flow, children need an appropriate word which would best represent their thoughts, and if they do not find the words in the input, they tend to create neologisms on their own, ascribing to them the regular grammatical morphemes.

The results of our analysis reveal the significant role of language input, and are in accordance with the findings of previous research, which show that these processes are universally present in early children's production in different languages. Furthermore, the process of semantic differentiation between verbal expressions seems to be a notable property of the morphologically rich languages such as Serbian and other Slavic languages, where derivation is a productive tool of new word formation.

Conclusion

The exploration of verbs in Serbian child language has revealed several patterns of verbal overgeneralizations that affect both the inflectional and derivational morphology. It has been shown that all parts of a verb could possibly be a subject to overgeneralizations: the stem, prefix, suffix, thematic vowel, inflection. The analysis has revealed the typical forms of Serbian overgeneralizations, represented by the forms **hoćem* 'want' 1.sg.pres., **nećem* 'want' 1.sg.pres.neg., **možem* 'can' 1.sg.pres. and **bide* 'be' 1.sg.pres. (inflectional and stem overgeneralizations). **Hoćem* and **nećem* have been detected as the developmental prerogatives among others, which bring an installation of overgeneralization as a resource for constructing the language specific verbal morphology.

A wide variety of facets has been found in the low frequent and hapax overgeneralizations recorded primarily in the sample of contentive verbs. Even though solitary and unrepeatably, their morphological properties exhibit the constraints of word formation in the Serbian language. Innovative solutions chosen by a child display both the regularities known to him/her and the features of grammar that are still missing in the system at the moment of the word composition. Since their occurrence is prolonged until the oldest age levels in the sample, the low frequent and hapax overgeneralizations reveal their significance in the growth of the vocabulary and adjustment of the new lexical items to the conjugational classes of the mother tongue.

The findings show that, although a statistically rare phenomenon, verbal overgeneralizations seem to be highly representative for the process of

Serbian language acquisition since they announce the onset of inflectional and derivational morphology in child language. Therefore, in monitoring the individual language development, overgeneralizations should not be treated as errors in production, but rather as typical and most creative parts of children's language construction.

References

- Authors (in preparation). Morphological properties of verbal overgeneralizations and neologisms in a sample of Serbian child language.
- Aguirre, C. (2003). Early Verb Development in one Spanish-speaking child. In D. Bittner, W.U. Dressler & M. Kilani-Schoch (Eds.), *Development of Verb Inflection in First Language Acquisition: A Cross-Linguistic Perspective* (pp. 1–25). Berlin: De Gruyter Mouton.
- Albright, A. & Hayes, B. (2003). Rules vs. analogy in English past tense: a computational/experimental study. *Cognition*, 90, 119–161.
- Ambridge, B., Freudenthal, D., Pine, J. M., Mills, R., Clark, V. & Rowland, C. F. (2009). Unlearning un-prefixation errors. In A. Howes, D. Peebles, & R. Cooper (Eds.), *9th International Conference on Cognitive Modelling* (pp. 158–163). Manchester: ICCM.
- Ambridge, B. (2010). Children's judgments of regular and irregular novel past-tense forms: New data on the English past-tense debate. *Developmental Psychology*, 46(6), 1497–1504. doi:10.1037/a0020668
- Ambridge, B., Pine, J. M., Rowland, C. F., Chang, F. & Bidgood, A. (2013). The retreat from overgeneralization in child language acquisition: Word learning, morphology, and verb argument structure. *Wiley Interdiscip Rev Cogn Sci* 4(1), 47–62. doi: 10.1002/wcs.1207
- Anderson, J. (1977). Induction of augmented transition networks. *Cognitive Science*, 1, 125–157.
- Anđel, M., Klampfer, S., Kilani-Schoch, Dressler, W. U. & Kovačević, M. (2000). Acquisition of verbs in Croatian, French and Austrian German – and outline of a comparative analysis. *Suvremena lingvistika*, 49–50, 5–25.
- Anđelković, D., Ševa, N. & Moskovljević, J. (2001). Serbian Corpus of Early Child Language. Laboratory for Experimental Psychology, Faculty of Philosophy, and Department of General Linguistics, Faculty of Philology, University of Belgrade. In *The CHILDES Database*. Available at: <http://childes.talkbank.org/access/Slavic/Serbian/SCECL.html>
- Arsenijević, B. (2006). *Inner aspect and telicity. The decompositional and the quantificational nature of eventualities at the syntax-semantics interface*. Utrecht: LOT Publications.
- Baker, C.L. (1979). Syntactic theory and the projection problem. *Linguistic Inquiry*, 10, 533–581.
- Baker, C. L. & MacCarthy, J. J. (Eds.) (1981). *The logical problem of language acquisition*. Cambridge: MIT Press.

- Bittner, D., Dressler, W. U. & Kilani-Schoch, M. (Eds.) (2003). *Development of Verb Inflection in First Language Acquisition: A Cross-Linguistic Perspective*. Berlin: De Gruyter Mouton.
- Bowerman, M. (1988). The 'no negative evidence' problem: How do children avoid constructing an overly general grammar? In J. Hawkins (Ed.), *Explaining language universals* (pp. 73–101). Oxford: Basil Blackwell.
- Brown, R. (1973). *A first language*. Cambridge, MA: Harvard University Press.
- Brooks, P., Tomasello, M., Dodson, K. & Lewis, L. (1999). Young overgeneralizations with fixed transitivity verbs. *Child Development*, 70, 1325–1337.
- Bybee, J. & Moder, C. L. (1983). Morphological classes as natural categories. *Language*, 59(2), 251–270.
- Clahsen, H. & Rothweiler, M. (1993). Inflectional rules in children's grammars: evidence from the development of participles in German. In G. Booij & J. van Marle (Eds.), *Yearbook of Morphology 1992* (pp. 1–34). Dordrecht: Kluwer Academic Publishers.
- Clahsen, H., Aveledo, F. & Roca, I. (2002). The development of regular and irregular verb inflection in Spanish child language. *Journal of Child Language*, 29, 591–622.
- Clark, E. (1983). Meanings and concepts. In J. Flavell, & E. Markman (Ed.), *Handbook of Child Psychology: Vol. 3, Cognitive Development* (pp. 787–840). New York: Wiley.
- Dressler, W. U. (1985). Typological aspects of Natural Morphology. *Wiener linguistische Gazette*, 36, 3–26.
- Dressler, W. U. (2005). Morphological typology and first language acquisition: some mutual challenges. In G. Booij, E. Guevara, A. Ralli, S. Sgroi, & S. Scalise (Eds.). *Morphology and linguistic typology. On-line Proceedings of the Fourth Mediterranean Morphology Meeting (MMM4), Catania, 21st –23rd September 2003* (pp. 7–20). Bologna: University of Bologna.
- Dressler, W. U., Dziubalska-Kořaczyk, K. & Katićić, A. (1996). A contrastive analysis of verbal inflection classes in Polish and Croatian. *Suvremena lingvistika*, 22(41–42), 127–138.
- Ervin, D. (1964). Imitation and structural change in children's language. In E. Lenneberg (Ed.), *New directions in the study of language* (pp. 163–189). Cambridge, Mass: MIT Press.
- Hržica, G. (2011). *Glagolske kategorije aspekta, vremena i akcionalnosti u usvajanju hrvatskog jezika*. Doktorski rad. Filozofski fakultet, Sveučilište u Zagrebu.
- Hržica, G. (2012). Daj mi to napisaj: preopćavanja glagolske osnove u usvajanju hrvatskog jezika. *Suvremena lingvistika*, 38(74), 189–208.
- Ilić, N. & Kostić, A. (2003a). *Podvrste homografije i automatska lematizacija teksta*. IX naučni skup Empirijska istraživanja u psihologiji [The 9th scientific conference Empirical research in psychology], Beograd, 13–14. februar 2003. *Rezimei*, str. 80–81.
- Ilić, N. & Kostić, A. (2003b). *Problem greške u automatskoj lematizaciji*. IX naučni skup Empirijska istraživanja u psihologiji [The 9th scientific conference Empirical research in psychology], Beograd, 13–14. februar 2003. *Rezimei*, str. 81–82.
- Li, P. & MacWhinney, B. (1996). Cryptotype, overgeneralization, and competition: A connectionist model of the learning of English reversive prefixes. *Connection Science*, 8(1), 3–30.

- Katičić, A. (2003). Early verb development in one Croatian speaking child. In B. Dągmar, W. U. Dressler & M. Kilani-Schoch (Eds.), *Development of verb inflection in first language acquisition* (pp. 239–267). Berlin: Mouton de Gruyter.
- Kostić, Đ. (1999). *Frekvencijski rečnik savremenog srpskog jezika*. [Frequency Dictionary of Contemporary Serbian Language]. Beograd: Institut za eksperimentalnu fonetiku i patologiju govora.
- Kuczaj, S. A. (1977). The acquisition of regular and irregular Past Tense forms. *Journal of Verbal Learning and Verbal Behavior*, 16, 598–600.
- MacWhinney, B. (1976). Hungarian research on the acquisition of morphology and syntax. *Journal of Child Language*, 3, 397–410.
- MacWhinney, B. (1988). Competition and teachability. In R. Schiefelbusch & M. Rice (Eds.) *The teachability of language* (pp. 63–104). New York: Cambridge University Press.
- MacWhinney, B. (1989). Competition and lexical categorization. In R. Corrigan, F. Eckman, & M. Noonan (Eds.), *Linguistic categorization* (pp. 195–242). New York: Benjamins.
- MacWhinney, B. (1993). The (il)logical problem of language acquisition. In *Proceeding of the 15th annual conference of the Cognitive Science Society* (pp. 61–70). Mahwah, NJ: Erlbaum.
- MacWhinney, B. & Snow, C. (1985). The Child Language Data Exchange System. *Journal of Child Language*, 12, 271–296.
- Mandić, M. (2013). Agreement morphology in early Serbian. In B. Surányi (ed.), *Proceedings of the Second Central European Conference in Linguistics for Postgraduate Students* (pp. 139–166). Budapest: Pázmány Péter Catholic University.
- Marcus, G., Pinker, S., Ullman, M., Hollander, M., Rosen, T. & Xu, F. (1992). Overgeneralization in language acquisition. *Monographs of the Society for Research in Child Development*, 57, 1–182.
- Nordquist, R. (2017). Overgeneralization Definition and Examples. *ThoughtCo. Glossary of Grammatical and Rhetorical Terms*. Humanities. Languages. Retrieved from: <https://www.thoughtco.com/overgeneralization-in-grammar-1691365>
- Perek, F. & Goldberg, A. E. (2015). Generalizing beyond the input: The functions of the constructions matter. *Journal of Memory and Language*, 84, 108–127. DOI: 10.1016/j.jml.2015.04.006
- Pinker, S. (1984). *Language learnability and language development*. Cambridge, Mass: Harvard University Press.
- Radisavljević, T. (2013). *Productivity in Serbian inflection and derivation in terms of Natural Morphology*. (Unpublished Master's Thesis). Universität Wien, Wien.
- Rumelhart, D. E. & MacClelland, J. L. (1986). On learning the past tenses of English verbs. In D.E. Rumelhart, J. L. MacClelland & The PDP Research Group (Eds.), *Parallel distributing processing: Explorations in the microstructure of cognition: Vol. 2* (pp. 216–271). Cambridge, MA: MIT Press.
- Skousen, R. (2001). *Analogy and Structure*. Dordrecht: Kluwer Academic.
- Slobin, D. (1973). Cognitive prerequisites for the development of grammar. In C. Ferguson & D. Slobin (Eds.), *Studies of child language development* (pp. 175–208). New York: Holt, Rinehart, & Winston.

- Stanojčić, Ž., Popović, Lj. (2000). *Gramatika srpskoga jezika*. [The grammar of Serbian]. Beograd: Zavod za udžbenike i nastavna sredstva.
- Stevanović, M. (1986). *Savremeni srpskohrvatski jezik I*. [The contemporary Serbo-Croatian I]. Beograd: Naučna knjiga.
- Theakston, A. (2004). The role of entrenchment in children's and adults' performance on grammaticality judgment tasks. *Cognitive Development*, 19, 15–24.

DATUM PRIJEMA RADA: 15.09.2017.

DATUM PRIHVATANJA RADA: 11.12.2017.

Distribucija hipergeneralizovanih oblika glagola u Srpskom elektronskom korpusu ranog dečijeg govora

Darinka Anđelković

*Laboratorija za eksperimentalnu psihologiju i Institut za psihologiju
Filozofski fakultet, Univerzitet u Beogradu*

Mirjana Mirić

Balkanološki institut, Srpska akademija nauka i umetnosti

Istraživanje je usmereno na pojavu hipergeneralizovanih oblika glagola snimljenih u uzorku spontane dečije produkcije na srpskom jeziku koji se odlikuje bogatom flektivnom i derivacionom morfologijom. *Srpski elektronski korpus ranog dečijeg govora* sastoji se od transkripata spontane jezičke produkcije osmoro dece longitudinalno snimanih od uzrasta 1;6 do 4;0. Pretraživani su hipergeneralizovani oblici glagola i istraživana je njihova distribucija i razvojna putanja kroz uzraste. Analiza je pokazala da su hipergeneralizovani oblici *hoćem 1.sg.prez., *nećem 1.sg.prez., *možem 1.sg.prez. i *bidem 1.sg.prez. najtipičniji predstavnici hipergeneralizacija u ranom dečijem govoru na srpskom jeziku. Oni se po pravilu registruju kod sve dece i to u relativno visokoj frekvenci. Hipergeneralizovano *hoćem i *nećem razvojno prethode svim drugim formama hipergeneralizacija glagola, ali se tokom razvoja ranije gube nego ostali oblici. Hipergeneralizacije nastale od punoznačnih glagola u dečijem govoru se pojavljuju nešto kasnije, obično su niskofrekventne ili hapaksi, ali se u detetovoj spontanoj produkciji zadržavaju dugo tokom razvoja. Nalazi su diskutovani i poređeni sa nalazima ranijih istraživanja, sa posebnim naglaskom na razvojne pravilnosti kroz uzraste.

Ključne reči: hipergeneralizacija glagolskih oblika, formiranje reči, korpus dečijeg govora, glagolska morfologija, srpski jezik