

UDK 811.163.42'242

811.163.42'367

Izvorni znanstveni rad

Rukopis primljen 14. IV. 2021.

Prihvaćen za tisak 16. XI. 2021.

doi.org/10.31724/rihjj.48.1.7

**Eva Pavlinušić Vilus**

Pomorski fakultet Sveučilišta u Rijeci

Studentska 2, HR-51000 Rijeka

orcid.org/0000-0002-8401-1270

*eva.pavlinusic@uniri.hr*

**Gordana Hržica**

Edukacijsko-rehabilitacijski fakultet Sveučilišta u Zagrebu

Borongajska cesta 83f, HR-10000 Zagreb

orcid.org/0000-0001-6067-9148

*gordana.hrzica@erf.unizg.hr*

## OMISSIONS AND OVERGENERALIZATIONS OF REFLEXIVE CLITIC IN THE ACQUISITION OF REFLEXIVE CONSTRUCTIONS IN CROATIAN AS L1

Verb constructions with a reflexive marker are termed “reflexive constructions”. Reflexive constructions in language acquisition research have been studied mainly within a formal theoretical framework while focusing on developmental differences in comprehension of syntactically bound or free pronominal elements. The present study aimed to determine the acquisition pathway of reflexive constructions in Croatian by examining the errors that children produce in early stages of acquisition. Correct and erroneous reflexive verb constructions were extracted from the spontaneous language production of three children previously recorded and transcribed for the Croatian Corpus of Child Language (Kovačević 2003). Errors were classified as omissions or overgeneralizations and further analyzed with respect to the type of reflexive construction, the complexity of the verb’s argument structure and consistency between the reflexive marker and verb. The results showed that children initially omitted the reflexive marker, then gradually introduced it into their production, occasionally overextending its use and thus producing overgeneralization errors. With age they became more successful in producing reflexive constructions. Consistent use of a reflexive marker alongside the verb in child-directed speech influenced the error rates in different types of reflexive constructions, while potential influence of the complexity of argument structure needs to be studied further.

## 1. Introduction

Verb constructions with a reflexive marker are termed ‘reflexive constructions’ (Geniušienė 1987). In some languages, reflexive markers are a specialized element, taking the form of an anaphor or a pronoun (Polish *się*, Croatian *se/sebe*), while in others a set of anaphoric or pronominal elements is used (English *myself*, *yourself* etc., French *me*, *te*, *se* etc.). There are also languages in which the reflexive marker is an affix, such as the Russian *-sja*. Despite the apparent morphological diversity of reflexive markers, the range of their meanings seems to be limited and constant across languages. The most common ones are reflexivity of action, reciprocity of action, decausative and inchoative meaning, middle, and passive/impersonal meaning (Geniušienė 1987; Marelj 2004).

Reflexive constructions (RCs) have been studied in Croatian only from the theoretical perspective (see, e.g., Belaj 2003; Oraić Rabušić 2015; Hrdlička 2020). In research on language acquisition in other languages, RCs have been studied mainly within the context of the “Delay of Principle B effect” or DPBE (Chien and Wexler 1990; Baauw 1999; Ruigendijk et al. 2010, among others): in various languages (e.g., English, Dutch, Hebrew), children interpret the personal pronoun in sentences such as *Lucie is washing her* as reflexive, i.e., syntactically bound. This indicates that children, unlike adults, allow the coreferential reading of a personal pronoun. The DPBE has been observed mainly during comprehension tasks and rarely during production, and it has not been reported in clitic languages (see, e.g., Zesiger et al. 2010 for French). Other aspects of RC acquisition have remained largely unexplored.

The aim of the present study was to examine RC acquisition in the Croatian language by focusing on errors that children make during acquisition.

### 1.1. RC acquisition

Reflexive markers are function words and so are not expected to be among the first words uttered by a child, since function words are generally acquired later than content words. Pronominal clitics are usually omitted in the early stages of language acquisition, and then gradually introduced into production (see, e.g.,

Pirvulescu et al. 2006 and Zesiger et al. 2010 for French, Costa and Lobo 2007 for Portuguese). A recent study by Varlokosta et al. (2016) on the acquisition of pronominal objects in non-clitic and clitic languages, including Croatian, showed that children in clitic languages master the usage of pronominal clitics by the age of five. However, Croatian children at that age still exhibit a small omission rate (6.4%, Varlokosta et al. 2016: 13). Omission of object clitics is not unexpected, since children omit verb arguments more frequently than adults do (Allen, Skarabela and Hughes 2008). Given structural and phonological similarities between reflexive and pronominal clitics, it seems reasonable to assume that acquisition of reflexive clitics follows a similar pattern as acquisition of object clitics. The work of Rivero and Goledzinowska (2002) on Polish children confirmed this to be the case. The authors examined the acquisition of constructions with the reflexive clitic *się*, which they found to be omitted entirely at the beginning of acquisition (Stage 1). The clitic made its first appearance in **intransitive** constructions by children between the ages 1;10 and 2;4 (Stage 2). Overgeneralizations with *się*, i.e. constructions in which the clitic was used with verbs that cannot form reflexive constructions, also appeared during Stage 2, as did RCs termed **intrinsically** reflexive, which the researchers analyzed as one-argument structures. [In this RC type, the reflexive marker can be omitted in some languages, e.g., in English: *I wash (myself) every morning* (Reinhart and Reuland 1993; Reuland 2003).] Stage 3 began with the appearance of **extrinsically** RCs (two-argument constructions with a reflexive marker). Thus, one-argument RCs preceded two-argument RCs in child language production, which is in accordance with Hale and Keyser's claim that multi-argument verb structure is derived via argument augmentation and verb transitivity from one-argument intransitive structure (Hale and Keyser 1993).

In their study of Polish children, Rivero and Goledzinowska (2002) examined only the role of argument structure complexity, yet other properties of RCs might also play a role during acquisition. One such property is how consistently the verb is marked as reflexive in child-directed speech, which may influence the child's subsequent production of omission and overgeneralization errors. In many languages, including Croatian, reflexive clitics accompany the verb *occasionally* as a marker of a specific semantic or syntactic process, or *consistently* as an obligatory morphological constituent of the verb. The latter group of RCs is known as *reflexiva tantum*. The consistency with which a linguistic ele-

ment appears in a certain structural context influences how easily the structure and element itself are acquired, as elaborated under the notion of cue reliability within the Competition Model (Bates and MacWhinney 1987). Cue reliability is a corpus measure that indicates how often the cue provides the correct interpretation. Since in reflexiva tantum the verb is marked as reflexive in all sentence contexts, it is a perfectly reliable cue for attaching the clitic *se*. Therefore, in the present study the role of cue reliability in acquisition of RCs in Croatian was examined, as well as the role of argument structure complexity.

## 1.2. RCs in Croatian

In Croatian, the clitic *se* takes the role of a reflexive marker and, as such, defines various properties of the verb in the RC, including semantic (reflexivity, reciprocity), syntactic (derived intransitivity) and pragmatic (agent-patient co-reference). While various schemes have been proposed to classify the argument structure of the verb and the role of the clitic in different types of RCs (for review see Oraić Rabušić 2015), the present work adopts the rather traditional classification into transitive or intransitive RCs (Belaj 2003; Silić and Pranjković 2005), allowing the results to be compared with those of Rivero and Goledzinowska (2002). Transitive RCs encompass constructions involving semantic reflexivity or reciprocity of action, as well as constructions termed “active-objectless”. In transitive RCs, *se* is analyzed as reflexive pronoun or anaphor playing the role of direct object. Intransitive RCs, in contrast, are subdivided into detransitivized constructions or reflexiva tantum, depending on the role of the clitic. In detransitivised RCs, the clitic is viewed as a marker of derived intransitivity of the verb, while the clitic in reflexiva Tantum is an obligatory morphological constituent of the verb. All transitive constructions with *se* and all detransitivized constructions contain verbs with a **transitive correlate**: the same verb can be used with or without *se*, unlike in reflexiva tantum.

The clitic *se* is also used to form middle, passive, and impersonal constructions. However, these constructions were not considered in the present study because they are analyzed as syntactic transformations, and our primary interest here was constructions in which *se* interacts with semantic, syntactic and/or pragmatic properties of the verb.

Table 1: Reflexive constructions in the Croatian language

REFLEXIVE CONSTRUCTION TYPE	Example	Role of the clitic <i>se</i>
<b>Transitive constructions with <i>se</i></b>		
<b>Semantically reflexive construction</b>	Ana            se            češlja. Ana.NOM.SG   se.REFL   comb.PRS.3SG 'Ana combs (herself).'	<i>se</i> = reflexive pronoun/anaphor, direct object
<b>Reciprocal construction</b>	Ivan            i    Ana            se            grle. Ivan.NOM.SG and Ana.NOM.SG se.REFL hug. PRS.3PL 'Ivan and Ana hug each other.'	<i>se</i> = reciprocal pronoun/anaphor, direct object
<b>Active-objectless construction</b>	Ivan            se            tuče. Ivan.NOM.SG   se.REFL   hit.PRS.3SG 'Ivan hits other people.'	<i>se</i> = indefinite pronoun?, direct object
<b>Intransitive constructions with <i>se</i></b>		
<b>Detransitivized construction</b>	Ana            se            zabavlja. Ana.NOM.SG   se.REFL   have fun.PRS.3SG 'Ana is having fun.'	<i>se</i> = detransitivization marker
<b>Reflexiva tantum</b>	Ana            se            smiješi. Ana.NOM.SG   se.REFL   smile.PRS.3SG 'Ana is smiling.'	<i>se</i> = morphological constituent

#### 1.4. Aim and hypotheses

Reflexive markers are acquired gradually, and their appearance depends on the type of construction. The formalist approach of Hale and Keyser (1993) predicts that children will produce more errors when producing transitive than intransitive RCs. The functionalist-oriented Competition Model (Bates and MacWhinney 1987) predicts that children will err more with inconsistently marked RCs than with consistently marked RCs, i.e., with reflexiva tantum. These two predictions overlap partially, as they both predict that children will produce more errors when producing transitive RCs than when producing reflexiva tantum.

Therefore, the goal of the present study was to examine errors that children make during RC acquisition in Croatian, while taking into consideration RC type in terms of verb argument structure and cue reliability.

We expected that:

H1. The error rate in RCs would diminish with age.

H2. Overgeneralizations of the clitic *se* would appear later than omissions of *se*.

H3. The rate of errors would differ across the three RC types as follows: the highest rate would be observed with transitive constructions (two-argument, low cue reliability), a lower rate with detransitivized constructions (one-argument, low cue reliability), and the lowest rate with reflexiva tantum (one argument, high cue reliability).

## 2. Methodology

### 2.1. Participants

Language samples of three children (Marina, Vjeran, Antonija) and their caretakers from the Croatian Corpus of Child Language (Kovačević 2003) were analyzed. Samples were taken two to three times per month from the onset of speech (Vjeran, Antonija) or slightly after that (Marina) until chronological age 2;8 (two children) and 3;2 (one child). A total of 136 samples was collected, comprising 42, 36 and 58 for the three children. Samples were prepared according to the Codes for the Human Analysis of Transcripts (CHAT) and were analyzed using the Computerized Language Analysis (CLAN) program (MacWhinney 2000).

### 2.2. Procedures

The CLAN program KWAL was used to extract utterances containing the clitic *se* from children's language production. In order to extract the omissions of *se*, all the verbs in children's utterances were first separated from all other word categories and then checked for the possibility of forming RCs. All the utter-

ances containing verbs identified as potential nuclei of RCs were extracted using CLAN software. Utterances that were unintelligible or that contained recitations, songs and poems were eliminated. The utterances in which *se* was used to form passive or impersonal constructions were not included in the analysis, since they constitute a separate group of constructions with *se*. Altogether 1067 RCs were extracted and stored in a spreadsheet, which contained the following data for every utterance: (1) chronological age (in months) when the utterance was produced, (2) the lemma of the verb used in the RC, (3) the morphological description of the form in which the verb lemma was used, (4) type of RC (transitive, detransitivized or reflexiva tantum), (5) correctness of utterance (correct, omission, overgeneralization) (following Rivero and Golezinska 2002), and (6) the number of tokens for the given RC.

The subcorpora of the three children differed in the number of recordings and in size. To minimize the effect of sample size differences, the basic unit of analysis was the percentage of utterances containing any RC or each type of RC per month of chronological age. We examined whether children erred more with certain RC types or with certain argument structures. The Kruskal-Wallis non-parametric test was used for group comparisons, while Spearman's non-parametric test was used to analyze potential correlations.

### 3. Results

#### 3.1. Rate of RC production by children

The percentages of the children's utterances that contained RCs were relatively low: Antonija, 2% (110 / 5649); Marina, 3% (317 / 9445); and Vjeran, 3% (644 / 20900). Nevertheless, the three children produced RCs nearly every month that they were recorded. Among the RC types, children most often produced detransitivized constructions, which accounted for 3.7% of Antonija's utterances per month, 4.6% of Marina's and 7.9% of Vjeran's.

### 3.2. Number and rate of errors

All three children erred in their production of RCs: more than 30% of RCs were erroneous in the case of Antonija and Vjeran, or around 15% in the case of Marina.

Table 2: Number of errors in reflexive constructions per month in child language

Participant	Minimum	Maximum	Mean	SD
Antonija	0	4	1.77	1.013
Marina	0	8	1.59	2.002
Vjeran	0	13	4.16	3.997

Table 3: Percentage of errors in reflexive constructions per month in child language

Participant	Minimum	Maximum	Mean	SD
Antonija	0	100	34.76	34.306
Marina	0	66.66	15.56	20.490
Vjeran	1.54	100	37.37	35.649

### 3.3. Types of errors

Children started producing RCs before they acquired the clitic *se*. This led to errors of omission and overgeneralization, which are illustrated below:

- a) Omission errors (0ref stands for the omission of reflexive clitic in the utterance)

#### Antonija, 1;7

Antonija: ovaj            0ref.            vuti.  
                   this.NOM.SG                    spin.PRS.3SG  
                   ‘This one /missing reflexive clitic/ spins’

Mother: ovaj            se            vrti.  
                   this.NOM.SG se.REFL            spin.PRS.3SG  
                   ‘This one spins.’



**Vjeran, 2;3**

Nanny:     sad     ideš             dolje?  
               now    go.PRS.2SG   down  
               ‘You are going down now?’

Vjeran:     ig(r)ati         0ref.  
               play.INF  
               ‘To play /missing reflexive clitic/’

b) Overgeneralization errors ([\*] stands for the overgeneralization of reflexive clitic in the utterance)

**Vjeran, 1;10**

Vjeran:     vidi             trčao             se [\*].  
               see.IMP.2SG   run.PTCP.M.SG   se.REFL  
               ‘Look, he was running /excessive usage of reflexive clitic/’

**Vjeran, 2;5**

Vjeran:     tata             se [\*]             zaspao.  
               dad.NOM.SG   se.REFL         fell\_asleep. PTCP.M.SG  
               ‘Dad /excessive usage of reflexive clitic/ fell asleep.’

The first RCs in which *se* is omitted appeared early: when Antonija was 1;7, when Marina was 1;5, and when Vjeran was 1;2. Omission errors persisted for quite a long time: 11 months in the case of Antonija, 13 in the case of Marina and until the end of recording in the case of Vjeran (1;2-3;2) (Figure 2). The first correct uses of *se* appeared 0-6 months after the first omissions, at respective ages of 1;10, 1;5 and 1;8. Interestingly, Marina at 1;5 simultaneously produced constructions with and without *se* (Figure 2). Unfortunately, this is when recording began for her, so it is unclear whether earlier she produced RCs lacking *se*.

Omissions of *se* and proper use of *se* appeared in transitive and detransitivized RCs before they appeared in reflexiva tantum. Overgeneralizations appeared at the ages of 2;4 (Antonija), 2;2 (Marina), and 1;10 (Vjeran), which was 8-9 months after the first omissions and 2-9 months after the first correct uses of *se*. The

children started overgeneralizing *se* with intransitive verbs (e.g., *ležati* ‘lie’, *puknuti* ‘snap’, *trčati* ‘run’), while overgeneralizations with transitive verbs appeared 0-3 months later and outnumbered overgeneralizations with intransitive verbs. Children produced overgeneralizations by using transitive verbs that cannot appear in RCs, such as *naučiti* ‘learn’, *kopati* ‘dig’, *svirati* ‘play (an instrument)’, *zaboraviti* ‘forget’, and *napraviti* ‘make’; and by using verbs that can form RCs but that were inappropriate for the sentence context because, for example, the sentence included a direct object NP. Examples of the latter types of verbs were *vidjeti* ‘see’, *igrati* ‘play’, *skinuti* ‘undress’, *prosuti* ‘spill’, and *zalijepiti* ‘glue’. In fact, Antonija produced only four utterances containing overgeneralizations of *se*; Marina, three utterances; and Vjeran, 17 utterances.

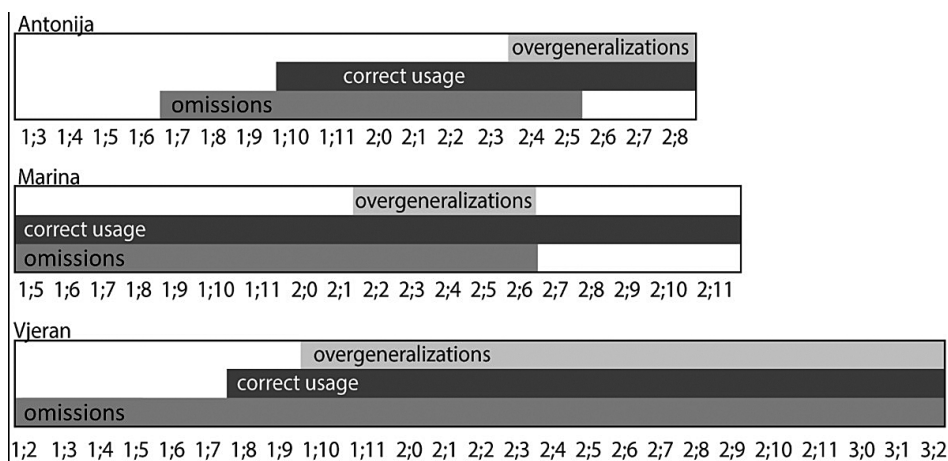


Figure 1: Usage of correct and erroneous reflexive constructions by children throughout the course of recording

The percentage of errors in child language decreased with chronological age ( $r_s = -.44, p < .01$ ), and this was true for each of the three children (Antonija:  $r_s = -.78, p < .01$ ; Marina:  $r_s = -.57, p < .05$ ; Vjeran:  $r_s = -.650, p < .05$ ).

### 3.4. Errors per RC type

Most RC errors were omission errors. Children committed these errors least often in reflexiva tantum RCs. Antonija and Vjeran erred more often with transitive RCs than detransitivized RCs, while Marina erred similarly often with the two types of RCs.

Table 4: Percentage of omission errors per reflexive construction type in child language

Participant	Error (%) per construction type	Minimum	Maximum	Mean	SD
Antonija	transitive	0	100	40.00	45.947
	detransitivized	0	100	23.50	36.706
	Reflexiva tantum	0	67	10.50	23.591
Marina	transitive	0	50	12.75	20.082
	detransitivized	0	83	17.41	24.285
	Reflexiva tantum	0	50	6.25	17.678
Vjeran	transitive	0	100	39.35	43.527
	detransitivized	0	100	25.52	30.688
	Reflexiva tantum	0	67	20.42	21.960

Across all RC types, all three children began making omission errors before or around the same time as they produced correct usage, and then subsequently they began making overgeneralizations. Two children began correctly using detransitivized verbs before they correctly used other RC types, whereas one child first produced transitive RCs correctly.

#### Antonija

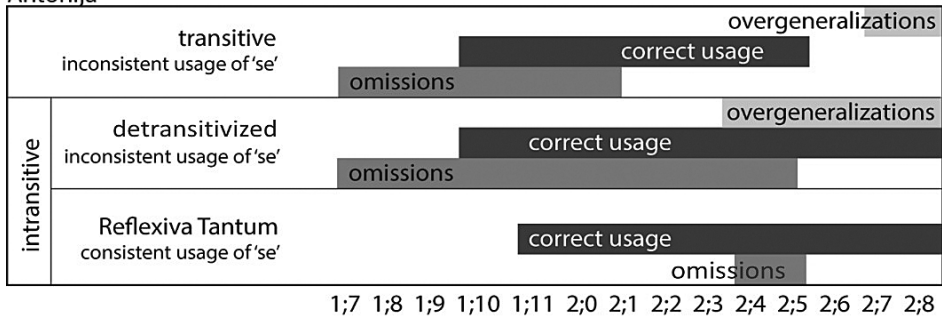


Figure 2: Correct and erroneous constructions in child language per type of reflexive construction throughout the course of recording (Antonija)

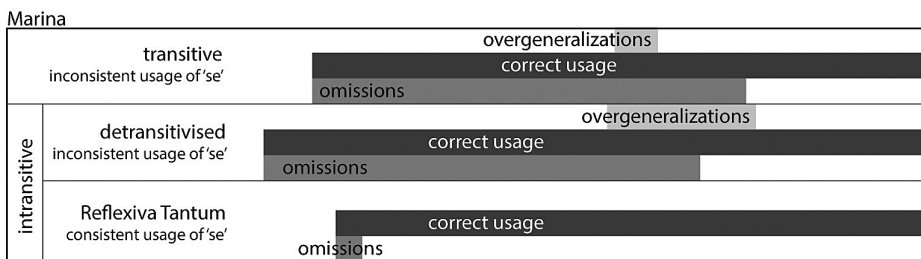


Figure 3: Correct and erroneous constructions in child language per type of reflexive construction throughout the course of recording (Marina)

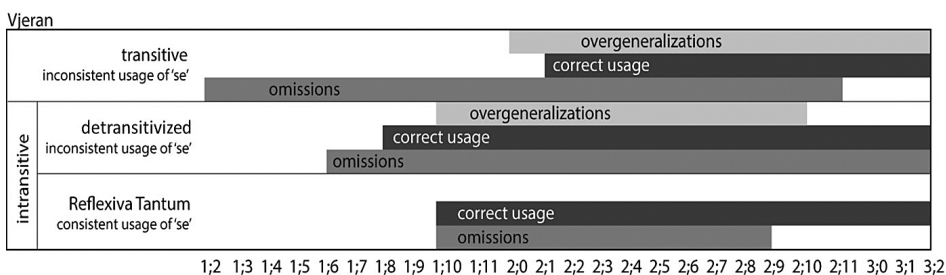


Figure 4: Correct and erroneous constructions in child language per type of reflexive construction throughout the course of recording (Vjeran)

Overall, the children erred most often with transitive RCs ( $M=48.41$ ,  $SD=43.075$ ), followed by detransitivized RCs ( $M=22.72$ ,  $SD=31.004$ ), and least often with reflexiva tantum ( $M=10.172$ ,  $SD=18.714$ ). The Kruskal-Wallis test showed these differences among RC types to be significant [ $H(2) = 16.57$ ,  $p<.001$ ]. Post-hoc Dunn-Bonferroni testing showed a significant difference between the error rates with transitive constructions or reflexiva tantum ( $p<.001$ ), and a marginally significant difference between the rates with detransitivized constructions or reflexiva tantum ( $p=.052$ ). The difference between the error rates with transitive or detransitivized constructions was not significant.

#### 4. Discussion

The results of the analysis corroborate the first hypothesis: for all three children, the percentage of RCs that were erroneous, because of either omission or

overgeneralization, decreased with age. Despite the complexity of the Croatian system of reflexive marking, children advance substantially in the correct usage of the clitic *se* by the age of three. Unfortunately, the recordings in the Croatian Corpus of Child Language do not extend beyond that age, so it remains unclear when children's usage of reflexive clitics approaches adult levels.

Consistent with previous studies, the children in our study produced errors first by omitting the clitic and later also by overextending its use during the first months of RC acquisition. The children here showed the same pattern of correct and incorrect usage as in the Polish children in the study by Rivero and Golezdzinowska (2002): complete omission of *se* > (occasional) correct usage of *se* > overgeneralization of the usage of *se*. These results corroborate the second hypothesis. Marina deviated from this pattern because she simultaneously started producing RCs with and without *se*. It is possible that she showed the same pattern as the other children, and her stage of exclusive omission of *se* preceded the start of her recording.

In all three children, earlier acquired forms of RCs coexisted with later ones: when children started adding *se* to the verb correctly, they still omitted the clitic in some RCs, and they began to make overgeneralizations with *se* before they stopped making omissions. Indeed, our data suggest that children pass through an acquisition stage during which correct RCs coexist with omissions and overgeneralizations.

Overgeneralizations with *se* appear in the speech of Croatian children towards the end of the second year of life, around the same age as in Polish children (Rivero and Golezdzinowska 2002). Those researchers define this as the second stage of RC acquisition. Functionalist approaches hold that overgeneralizations of linguistic structures reflect the transition from knowledge and usage based on single items, towards an understanding of underlying structural patterns and their application (Brooks et al. 1999). Whereas Rivero and Golezdzinowska (2002) reported overgeneralizations only with intransitive verbs, the children in the present study overextended reflexive markers with transitive verbs as well. Nevertheless, overgeneralizations of *se* seemed to appear earlier with intransitive verbs (see Figures 3-5). These results likely reflect that as children are exposed to transitive verbs that can be reflexive or non-reflexive, they start adding

*se* to transitive verbs that are never used in RCs or to verbs that can be used in RCs, but not in the given sentence context.

Analysis of the rate of errors in the three RC types corroborated the third hypothesis: the rate of omission errors was highest with transitive RCs, followed closely by the rate of omission errors with detransitivized RCs, and lowest with reflexiva tantum. Error rates differed only slightly for different RC types, except for the difference between transitive RCs and reflexiva tantum. Given the small size of the sample, this is not unexpected. As predicted, omission errors were more frequent with two-argument RCs than with one-argument RCs, which is in accord with formalistic theory about the role of argument structure complexity in the acquisition of verbs (Hale and Kayser 1993). However, functionalist studies on verb acquisition have also suggested that other reasons may lead children to start producing intransitive verbs and verb frames earlier than transitive ones. For example, intransitive verbs and verb frames may occur more frequently in child-directed speech, which would mean that children are more exposed to verbs with one-argument structure and therefore start producing them earlier than two-argument verbs (Theakston et al. 2001). Preliminary research on our corpus of three children indicates that intransitive RCs may be more frequent than transitive RCs in the child-directed speech in the recording, which may help explain why children err less with that RC type. This issue requires further study, since the analysis of the distribution of different RC types in the input is beyond the scope of the present work.

Our analysis also demonstrated that omission errors were more frequent in RCs in which the verb was inconsistently reflexive-marked than in RCs in which the verb was always accompanied by *se* (reflexiva tantum). This result supports the prediction based on the notion of cue reliability (Bates and MacWhinney 1987). Verbs from the reflexiva tantum group always appear with *se* in the input, which makes them perfectly reliable cues for the usage of the clitic, and children seem to be sensitive to this regularity. Also in accord with cue reliability, all three children in our study omitted *se* from reflexiva tantum later than they began omitting it from the other two RC types. The gap between the first omission of *se* from transitive or detransitivized RCs and subsequent first omission of *se* from reflexiva tantum was 8 months for Antonija and Vjeran and 2 months for Marina. Interestingly, omissions with reflexiva tantum appeared simultaneously

with, or 0-4 months after, the first correct uses of this RC type, which was not the case for the other two RC types. In fact, omissions of *se* in reflexiva tantum appeared almost simultaneously with its overgeneralizations but not with its omissions in other RC types (see Figures 3-5).

Our results suggest that at the time of the first omission errors, children are in a stage of complete omission of function words. The appearance of the first correct RCs with *se* fits well into the stage of conservative production described within Tomasello's (2000) "verb-island hypothesis". Children try to produce the verbs as they hear them in the input and start attaching clitics to them. In the case of verbs that appear in transitive and detransitivized RCs, which can be transitive (without *se*) or reflexive (with *se*) in the input, children must rely on their memory to produce these verbs with or without *se* in reflexive and non-reflexive contexts, with more or less success. During this stage, children are reluctant to omit *se* in reflexiva tantum because they never hear these verbs without *se*. Only upon entering the stage of pattern recognition and generalization do children start omitting the clitic even in reflexiva tantum, because of confusion about the variability of RCs with or without *se* in the input. The appearance of this type of *se* omission as well as *se* overgeneralization marks the end of conservative, item-based usage of verb constructions.

This research has limitations. First, it was conducted on a relatively small corpus (around 90 000 tokens of child speech). Since RCs are not frequent, this led to a modest number of expressions to analyze. The other limitation lies in the sampling method of the Croatian Corpus of Child Language, which features a relatively small number of child recordings per month. The probability of capturing an error in speech depends on the frequency of the error and the sampling density (for an overview, see Rowland, Fletcher and Freudenthal 2008). Sampling similar to the one in our corpus may not be sufficient to capture low- or even medium-frequency errors. Future research might implement different sampling regimes (including cross-sectional sampling) or different research methods to yield more abundant data.

## 5. Conclusion

Analysis of the acquisition of RCs in Croatian has shown that children follow the general pattern described by Rivero and Golezinska (2002) for Polish: they start producing RCs with *se* omitted, then gradually produce more and more utterances with *se* used correctly, with occasional overgeneralization of the clitic. With time, the error rates in children's production of RCs decrease. Cue reliability influences the acquisition of RCs, while the role of the argument structure of the verb remains to be further studied.

To the best of our knowledge, this study represents the first research of the early acquisition of RCs in Croatian. It is also one of the few studies performed in a language with a complex system of RCs that includes reflexiva tantum. Since languages differ in morphological expression of reflexive markers, the present work adds to the body of cross-linguistic research aiming to describe the acquisition of reflexive constructions.

## Acknowledgments

This work has been supported in part by the Croatian Science Foundation under the project English words in Croatian: Identification, affective-semantic norming and investigation into cognitive processing via behavioural and neuroscientific methods (UIP-2019-04-1576) and under the project Multilevel approach to spoken discourse in language development (UIP-2017-05-6603).

## References

- ALLEN, SHANLEY; SKARABELA, BARBORA; HUGHES, MARY. 2008. Using corpora to examine discourse effects in syntax. *Corpora in language acquisition research: Finding structure in data* 99–137. doi.org/10.1075/tilar.6.07all.
- BATES, ELISABETH; MACWHINNEY, BRIAN. 1987. Competition, variation, and language learning. *Mechanisms of language acquisition*. Ed. MacWhinney, Brian. Lawrence Erlbaum Associates. Hillsdale, NJ. 157–194.
- BAAUW, SERGIO. 1999. The Role of the Clitic–Full Pronoun Distinction in the Acquisition of Pronominal Coreference. *Proceedings of the 23rd Annual Boston University Con-*



- ference on Language Development*. Ed. Greenhill, Anabel; Littlefield, Harriet; Tano, Cheryl. Cascadilla Press. Sommersville. 32–43.
- BELAJ, BRANIMIR. 2001. Prototipno-kontekstualna analiza povratnih glagola u hrvatskom jeziku. *Suvremena lingvistika* 51. 1–11.
- BROOKS, PATRICIA J. ET AL. 1999. Young children's overgeneralizations with fixed transitivity verbs. *Child Development* 70. 1325–1337. doi.org/10.1111/1467-8624.00097.
- CHIEN, YU-CHIN; WEXLER, KENNETH. 1990. Children's Knowledge of Locality Conditions in Binding as Evidence for the Modularity of Syntax and Pragmatics. *Language Acquisition* 1. 225–295. doi.org/10.1207/s15327817la0103\_2.
- COSTA, JOÃO; LOBO, MARIA. 2007. Clitic omission, null objects or both in the acquisition of European Portuguese. *Romance Languages and Linguistic Theory*. Eds. Baauw, Sergio; Drijknonongen, Frank; Pinto, Manuela. John Benjamins. Amsterdam. 59–72. doi.org/10.1075/cilt.291.06cos.
- FEHRMANN, DOROTHEE; JUNGHANNS, UWE; LENERTOVÁ, DENISA. 2010. Two reflexive markers in Slavic. *Russian Linguistics* 34. 203–238. doi.org/10.1007/s11185-010-9062-7.
- GENUŠIENĖ, EMMA. 1987. *The Typology of Reflexives*. Mouton de Gruyter. Berlin. doi.org/10.1515/9783110859119.
- HALE, KEN; KEYSER, SAMUEL J. 1993. On Argument Structure and the Lexical Expression of *Syntactic Relations*. *The View from Building 20: Essays in Linguistics in Honor of S. Bromberger*. Eds. Hale, Ken; Keyser, Samuel J. MIT Press. Cambridge, MA.
- HRDLIČKA, MIROSLAV. 2020. *Kategorija povratnosti u hrvatskom i poljskom jeziku*. Doctoral dissertation. University of Zagreb. Zagreb. 211 pages.
- KOVAČEVIĆ, MELITA. 2003. Croatian Corpus. <https://childes.talkbank.org/access/Slavic/Croatian/Kovacevic.html>.
- MACWHINNEY, BRIAN. 2000. *The CHILDES Project: tools for analyzing talk*. Erlbaum. Hillsdale. doi.org/10.1177/026565909200800211.
- MARELI, MARIJANA. 2004. *Middles and argument structure across languages*. LOT. Utrecht.
- ORAIĆ RABUŠIĆ, IVANA. 2015. O odnosu elementa *se* i zamjenice *sebe*. *Rasprave Instituta za hrvatski jezik i jezikoslovlje* 41/1. 97–126.
- PIRVULESCU, MIHAELA. 2006. Theoretical Implications of Object Clitic Omission in Early French: Spontaneous vs. Elicited Production. *Catalan Journal of Linguistics* 5. 221–236. doi.org/10.5565/rev/catjl.70.
- REINHART, TANYA; REULAND, ERIC. 1993. Reflexivity. *Linguistic Inquiry* 24. 657–720.
- REULAND, ERIC. 2003. Anaphoric dependencies: A window into the architecture of the language system. *Glott International* 7/1–2. 3–25.
- RIVERO, MARÍA L.; GOLEDZINOWSKA, MAGDALENA. 2002. The acquisition of constructions

with reflexive clitics in Polish. *Asymmetry in Grammar: Volume 2: Morphology, phonology, acquisition*. Ed. Di Sciullo, Anna M. John Benjamins, Amsterdam. 265–289. doi.org/10.1075/la.58.11riv.

ROWLAND, CAROLINE F.; FLETCHER, SARAH L.; FREUDENTHAL, DANIEL. 2008. How big is big enough? Assessing the reliability of data from naturalistic samples. *Corpora in Language Acquisition Research: History, methods, perspectives*. Ed. Behrens, Heike. John Benjamins. 1–24. Amsterdam/Philadelphia. doi.org/10.1075/tilar.6.04row.

RUIGENDIJK, ESTHER ET AL. 2010. Symmetry in comprehension and production of pronouns: A comparison of German and Hebrew. *Lingua* 120/8. 1991–2005. doi.org/10.1016/j.lingua.2010.02.009.

SILIĆ, JOSIP; PRANJKOVIĆ, IVO. 2005. *Gramatika hrvatskoga jezika: za gimnazije i visoka učilišta*. Školska knjiga. Zagreb.

TOMASELLO, MICHAEL. 2000. Do young children have adult syntactic competence? *Cognition* 74. 209–253. doi.org/10.1016/S0010-0277(99)00069-4.

VARLOKOSTA, SPYRIDOULA ET AL. 2016. A cross-linguistic study of the acquisition of clitic and pronoun production. *Language Acquisition* 23/1. 1–26. doi.org/10.1080/10489223.2015.1028628.

ZESIGER, PASCAL E. ET AL. 2010. The acquisition of pronouns by French children: A parallel study of production and comprehension. *Applied Psycholinguistics* 31. 571–603. doi.org/10.1017/S0142716410000147.

## **Ispuštanja i poopćavanja povratne zamjenice u usvajanju povratnih konstrukcija u hrvatskom kao J1**

### *Sažetak*

Glagolske konstrukcije u kojima se uz glagol nalazi povratna oznaka nazivaju se povratnim konstrukcijama (eng. *reflexive constructions*). Većina istraživanja usvajanja povratnih konstrukcija provedena je na engleskom govornom području u okviru formalističkih pristupa usvajanju jezika, dok su istraživanja u ostalim jezicima malobrojna, unatoč činjenici da je u mnogim jezicima uloga povratne oznake složenija nego u engleskome. Cilj istraživanja predstavljenog u ovom radu bio je opisati tijek ranog usvajanja povratnih konstrukcija u hrvatskome kao prvom jeziku kroz analizu jezičnih pogrešaka djece. Iz zapisa spontane jezične proizvodnje troje djece od progovaranja do 3. godine (Croatian Corpus of Child Language, Kovačević 2003) izdvojeni su iskazi koji sadrže povratne konstrukcije u ispravnom ili neispravnom obliku. Potonji su analizirani u odnosu na vrstu pogreške – izostavljanje povratne oznake *se* ili poopćavanje njezine uporabe, složenost argumentne strukture glagola te dosljednost uporabe povratne ozna-

ke uz glagol u povratnoj konstrukciji u ciljnom jeziku. Rezultati su pokazali da djeca započinju proizvodnju povratnih konstrukcija u potpunosti izostavljajući povratnu oznaku koju s vremenom počinju rabiti uz glagol, ponekad i u konstrukcijama u kojima njezina uporaba nije ispravna, čime nastaju pogreške poopćavanja uporabe nenaglasnice *se*. Udio pogrešaka u proizvodnji povratnih konstrukcija smanjuje se s povećanjem dobi djece. Dosljednost uporabe povratne oznake uz glagol u ciljnom jeziku pokazala se kao značajni čimbenik u pogrešnoj proizvodnji glagolskih konstrukcija, dok utjecaj složenosti argumentne strukture glagola treba još istražiti.

**Keywords:** corpus study, reflexive constructions, reflexive marker, clitic *se*, omission errors, overgeneralization errors

**Ključne riječi:** korpusno istraživanje, povratne konstrukcije, povratna oznaka, nenaglasnica *se*, pogreške izostavljanja, pogreške poopćavanja

