volume 27, number 1

winter–spring 2019

Journal of the Slavic Linguistics Society

Slavic Linguistics Society

The Slavic Linguistics Society is an organization devoted to the systematic and scholarly study of the Slavic languages. Membership in the Society is open to all linguists regardless of field, theory, framework, or school. Members are entitled to present papers at annual meetings and receive the *Journal of Slavic Linguistics*, which is the official journal of the Slavic Linguistics Society. Individuals receive a subscription to *JSL* with their annual membership in the Slavic Linguistics Society. To join, go to http://www.slaviclinguistics.org.

SLS is governed by an Executive Board consisting of six members elected at the annual meeting for three-year staggered terms. The Executive Board negotiates arrangements for the annual meetings with host institutions. Meetings alternate between North America and Europe.

Executive Board

Guidelines for Contributors

JSL is intended to address issues in the description and analysis of Slavic languages of general interest to linguists, regardless of theoretical orientation. Papers dealing with any aspect of synchronic or diachronic Slavic phonetics, phonology, morphology, syntax, semantics, or pragmatics will be considered, as long as they raise substantive problems of broad theoretical concern or propose significant descriptive generalizations. Comparative studies and formal analyses are especially encouraged. Submissions ranging from full-length articles to shorter remarks and replies will be considered, along with substantive reviews.

Manuscripts should be submitted through the journal's submission website http://ojs.ung.si/index.php/JSL in .pdf format. All submissions should follow the submission guidlines published on the journal website. Papers should be anonymous and include a one-paragraph abstract. A computer file (preferably .doc[x]) will be requested before publication; we ask that you use only Unicode fonts.

Visit our World Wide Web page: <http://ojs.ung.si/index.php/JSL>

The *Journal of Slavic Linguistics* is published twice annually. Each volume consists of two issues, totaling approximately 360 pages: Winter–Spring and Summer–Fall. Library and institutional subscriptions are available from Slavica Publishers at the following rate:

Institutions $60.00

Individuals should join the Slavic Linguistic Society (http://www.slaviclinguistics.org), which carries with it an annual subscription to *JSL*.

Airmail supplement for institutions outside of North America: $11.00/year.

A limited number of back issues are available for $24.00 apiece.

Indexing and Abstracting: *American Bibliography of Slavic and East European Studies, ERIH (European Reference Index for the Humanities and Social Sciences), Humanities International Index, IBZ (Internationale Bibliographie der Geistes- und Sozialwissenschaftlichen Zeitschriftenliteratur), MLA International Bibliography (Modern Language Association), OCLC ArticleFirst, and Web of Science Emerging Sources Citation Index*

This issue was published in October 2019.

ISSN: 1068-2090

Technical Editor: Renata Uzzell

Cover design: Nikki Life

Slavica Publishers
Indiana University
1430 N. Willis Drive
Bloomington, Indiana USA
47404-2146

[Tel.] 1-812-856-4186
[Fax] 1-812-856-4187
[Toll-free] 1-877-SLAVICA
slavica@indiana.edu
http://www.slavica.com

journal of slavic linguistics

volume 27, number 1

winter–spring 2019

articles

reviews

North Slavic *-ě* vs. South Slavic *-ę*: A Problem of Forward Reconstruction*†

Ronald I. Kim

Zum Andenken an Charlie Townsend
Na památku Charlieho Townsenda
Ku pamięci Charliego Townsenda
To the memory of Charles E. Townsend

Abstract: The long problematic correspondence of North Slavic *-ě* ~ South Slavic *-ę* in the *jo*-stem accusative plural and *jā*-stem genitive singular and nominative and accusative plural is best explained by positing a Proto-Slavic contrast within the soft *jā*-stems between gen. sg., nom. pl. **-ē* and acc. pl. **-ę̄*, which was leveled in different directions as NSl *-ě* and SSl *-ę*. With its nasal vowel, the acc. pl. ending must go back to **-jāns*, thereby demonstrating that the PIE *eh*$_2$-stem desinence **-eh$_2$s* > **-ās* was remade to **-āns* after the other declensional classes in the early prehistory of Slavic. The Baltic facts are consistent with a Proto-Balto-Slavic date for this innovation.

* More than anyone else, Charlie Townsend was responsible for kindling my lifelong interest in Slavic linguistics. Even now, more than two decades after taking his graduate course in Common and Comparative Slavic, I continually refer in my everyday experiences with Slavic languages to the facts and ideas I first encountered there, as well as his many amusing anecdotes about language learning and travel in the Slavic world.

† Versions of this paper and its companion (Kim forthcoming b) were presented at the 36th East Coast Indo-European Conference held at Cornell University on 1–4 June 2017; at the Seminar für Indogermanistik, Friedrich-Schiller-Universität Jena on 6 Nov. 2017; and at the Instytut Językoznawstwa, Uniwersytet Jagielloński on 27 Nov. 2017. I thank the participants on those occasions for their remarks and suggestions, in particular Yaroslav Gorbachov, Jay Jasanoff, Martin Kümmel, Craig Melchert, Sergio Neri, Dariusz Piwowarczyk, Don Ringe, Wojciech Smoczyński, and Michael Weiss, as well as Joseph Eska, Reiner Lipp, Marek Majer, and two anonymous reviewers for their valuable comments.

Additional abbreviations: Av = Avestan; ill. = illative; InIr = Indo-Iranian; NSl = North Slavic; ONovg = Old Novgorodian; OPr = Old Prussian; Ved = Vedic.

This article addresses a problem of nominal inflection which has long been noted in Slavic comparative grammars but for which scholarly opinion is arguably no closer to a consensus today than a century ago. This is the so-called *ě tertium* or **ě*$_3$, namely, the correspondence North Slavic *-ě* : South Slavic *-ę* in the accusative plural of *jo*-stems and the genitive singular, nominative plural, and accusative plural of *jā*-stems. Section 1 introduces the inflection of PSl *(j)o-* and *(j)ā*-stems and the sound changes involved in their development from PIE. Section 2 then reviews the long history of research on the problem of North Slavic *-ě* vs. South Slavic *-ę* and concludes that none of the phonological or analogical explanations proposed so far is satisfactory. Progress towards a solution can be achieved if one takes into full account the evidence of the Baltic and other Indo-European languages and employs the method of forward reconstruction from Proto-Indo-European to Proto-Slavic. It is argued that the PIE *eh*$_2$-stem accusative plural **-eh*$_2$*s* > **-ās* was remade to **-āNs* after the other stem classes and that this is the source of the nasal vowel in South Slavic, where the ending *-ę* was generalized to the gen. sg. and nom. pl. forms; conversely, the northern dialects generalized gen. sg. and nom. pl. *-ě* to the acc. pl. of both *jā-* and *jo*-stems (section 3). The Baltic facts suggest, if they do not prove, that the remodeling of the *ā*-stem acc. pl. ending to **-āns* took place already in Proto-Balto-Slavic (section 4).

1. Introduction: Slavic *(j)o-* and *(j)ā*-stems

Slavic historical grammar traditionally distinguishes between masculine *o-* and *jo*-stems, which continue respectively PIE nonneuter stems in **-o-* and **-yo-*, and similarly between *ā-* and *jā*-stems, which continue respectively (post-)PIE stems in **-eh*$_2$*-* and **-yeh*$_2$*-*. These inflectional types are illustrated in Table 1 on the following page with PSl **gardu* 'enclosure, fort, town', **mąži* 'man, husband' (OCS *gradŭ*, *mǫžĭ*) and **ženā* 'woman, wife', **dūšā* 'soul' (OCS *žena*, *duša*), which continue respectively pre-PSl **gardas*, **mangjas* and **ženā*, **dauxjā*.[1]

[1] The PSl reconstructions adopted here are those which in my view may be arrived at by applying the Comparative Method to the attested medieval and modern Slavic languages, with minimal reference to data from Baltic or other Indo-European languages. They are thus intermediate between the "Early Common Slavic" and "Late Common Slavic" of Shevelov (1964), Birnbaum (e.g., 1998), and others, though closer to the latter, and not nearly as radical as those of Holzer (e.g., 1998, 2003), who takes into account all the onomastic evidence from neighboring non-Slavic languages of the 1st millennium AD (see Kim forthcoming a).

The literature on PSl nominal inflection is enormous; for an extremely useful overview, see Olander 2015. Note that the ending of the *jo*-stem acc. pl. and *jā*-stem gen. sg., nom. pl., and acc. pl. is provisionally given in the table as **-ē* (OCS *-ę*).

Table 1

nom. sg.	*gardu	*mąži	*ženā	*dūšā[2]
gen.	*gardā	*mąžā	*ženȳ	*dūšę̄
dat.	*gardū	*mąžū	*ženē	*dūšī
acc.	*gardu	*mąži (*mąžā)	*ženą	*dūšą
inst.	*gardami	*mąžemi	*ženajā	*dūšejā
loc.	*gardē	*mąžī	*ženē	*dūšī
voc.	*garde	*mąžū	*žena	*dūše
nom./acc. du.	*gardā	*mąžā	*ženē	*dūšī
gen./loc.	*gardū	*mąžū	*ženū	*dūšū
dat./inst.	*gardamā	*mąžemā	*ženāmā	*dūšāmā
nom./voc. pl.	*gardī	*mąžī	*ženȳ	*dūšę̄
gen.	*gardu	*mąži	*ženu	*dūši
dat.	*gardamu	*mąžemu	*ženāmu	*dūšāmu
acc.	*gardȳ	*mąžę̄	*ženȳ	*dūšę̄
inst.	*gardȳ	*mąžī	*ženāmī	*dūšāmī
loc.	*gardēxu	*mąžīxu	*ženāxu	*dūšāxu

The difference between the "hard" *o*- and *ā*-stem declensions and the "soft" *jo*- and *jā*-stem declensions is the result of the pre-PSl fronting of back vowels after **j* (i.e., **ja* > **je* and **ju* > **ji*), and various other conditioned sound changes, many specific to word-final position. As the Slavic *Auslautgesetze* remain hotly disputed and can hardly be examined here in detail, the remainder of this section will briefly present the views assumed in this study, to set the background for the discussion of the problematic ending *-ę ~ -ě* in sections 2–3.

The fronting of **ja* > **je* was itself preceded by the raising of **a* > **u* under certain conditions in final syllables, although the exact conditions for this raising are notoriously controversial, above all the phonologically regular reflex of word-final PIE **-os* > pre-PSl **-as*.[3] According to the general consensus, pre-PSl **a* > **u* before a nasal, as in the accusative singular of masculine *o*-stems and the first person singular of thematic aorists (Table 2 on the following page).

[2] For the reconstruction of *jā*-stem nom. sg. **-ā*, dat./inst. du. **-āmā*, dat. pl. **-āmu*, inst. pl. **-āmī*, loc. pl. **-āsu* and *jo*-stem gen. sg. **-ā*, nom./acc. du. **-ā*, see fn. 25 below.

[3] Contrast the recent treatments of Vermeer 1991; Majer 2011; Kortlandt 2016: 91 (PIE **-os* > PSl **-o*, i.e., **-a*) and Olander 2012, 2015: 56 (PIE **-os* > PSl **-ə* > ONovg *-e*, elsewhere *-ŭ*). For the voluminous literature on this question, see Majer 2011: 352–53 and Olander 2012: 321–26, 2015: 102–4.

Table 2

	PIE		PBSl			PSl
o-stem masc. acc. sg.	*-om	>	*-an	>	*-un	> *-u (OCS -ŭ)[4]
thematic aor. 1sg.	*-om	>	*-an	>	*-un	> *-u (OCS -ŭ)

From the thematic aorist (e.g., OCS *padŭ* 'I fell'), the ending **-u(n)* spread to the sigmatic aorist, replacing the expected reflex of PIE **-ṇ* (e.g., PIE **rḗk-s-ṇ* > **rēksin* > **rēxi(n)* → PSl **rēx-u* > OCS *rěxŭ* 'I said'), as well as the specifically Slavic extended aorist (e.g., PIE **ĝneh₃-ṇ* → PSl **znā-x-u* > OCS *znaxŭ* 'I knew').[5] The 1sg. nom. personal pronoun also shows this development: PIE **eĝh₂-om* (Ved *ahám*, Av *azəm*) > PSl **ēzu* > OCS *azŭ*, Blg *az*, Sln *jaz*. The gen. pl. may as well, though its reconstruction remains controversial: if the PIE ending was **-Hom*, this would have become **-on* > **-an* > **-un* > PSl **-u* (OCS *-ŭ*) and could subsequently have been extended to the *o*- and *ā*-stems.[6] On the other hand, raising apparently did not affect pre-PSl **ā* (< post-PIE **ā*, **ō*), on the evidence of PIE *ā*-stem acc. sg. **-eh₂m* > PBSl **-ān* > PSl **-ą̄* (OCS *-ǫ*).[7]

There was also a separate raising of pre-PSl **ā* > **ū* before word-final **-(n)s*, to judge by the ending of the masculine *o*-stem accusative plural and of the *ā*-stem genitive singular and nominative plural (Table 3 on the following page).[8]

[4] Orr (1985: 23–25, 1986, 1988, 2000: 134–37, 142–45) argues against the whole notion of *Auslautgesetze* and thinks that the regular reflex of PIE **-om* was **-ǫ*, but this necessitates massive interparadigmatic analogy to explain the forms discussed here. Note also that the reconstruction of PIE acc. sg. **-o-m* vs. **-u-Ø*, **-i-Ø* (Orr 2000: 55–69) finds absolutely no support in the comparative evidence. Kortlandt (1978: 286–90, 1983: 173–74, 2016: 88–89) assigns the raising **-om* > **-um* to PBSl on the basis of Old Prussian forms in *-on* as well as accentual considerations, which cannot be entered into here.

[5] See Olander 2015: 310, with references to alternative views.

[6] See Olander 2015: 255–57, with references to alternative views. On the reconstruction of the PIE genitive plural ending, see among others Kortlandt 1978 (**-om*); Ringe 2006: 41, 2017: 50, 51 (**-oHom*, thematic **-o-oHom*); Kümmel 2013 (PInIr **-ăHăm* < PIE **-oHom* or **-eHom*). Jasanoff (2014; 2017: 152) proposes instead that PIE **-oHom* > **-ōn* > **-ūn*, with the same raising of nonacute **ō* > **ū* as in *n*-stem nom. sg. **-ō* > **-ū* > PSl **-ȳ* (OCS *kamy* 'stone', Pol *kamy-k* 'pebble'; see Jasanoff 1983: 141, 146–49, 2017: 57–58, and already Streitberg 1891: 294–96); this in turn gave a unique segment **-ụ*, reinterpreted as PSl **-u* with metatony of the preceding syllable peak.

[7] Cf. Viredaz 2009: 9 and Olander 2015: 116–18 with refs.; pace Matasović 2008: 123–24, 192.

[8] First proposed by Hirt (1893: 353–54), followed by Meillet (1897: 96, 104, 125–26, 1902: 109, 1914–15: 6–7, 1934: 151, 398), Vondrák (1898: 337, 338–39, 1904: 187, 1906: 52–53, 108; later abandoned, cf. Vondrák 1924: 151–52), Lehr (1917: 39), Milewski (1932: 27–29), and

Table 3

	PIE	PBSl				PSl
o-stem masc. acc. pl.[9]	*-ōms	> *-ōns	> *-āns	> *-ūns	> *-ūs	> *-ȳ (OCS -y)
ā-stem gen. sg.	*-eh$_2$es	> *-ās	>	*-ūs	>	*-ȳ (OCS -y)
ā-stem nom. pl.	*-eh$_2$es	> *-ās	>	*-ūs	>	*-ȳ (OCS -y)

Post-PIE *-*ās* in the *ā*-stem gen. sg. and nom. pl. thus would **not** have become PSl *-*ā* and merged with the nom. sg., and there is no need to ascribe the ending *-*ȳ* to analogical extension from the acc. pl., as assumed by most scholars ever since Scherer (1868: 291, 474–75).[10] The other popular explanation of OCS *ā*-stem gen. sg. *-y* as a transfer from nasal-stem inflection, usually *-*ŏ̄n-s vel*

Galabov (1973: 10); for other references, see Hujer 1910: 100 and Olander 2012: 332, fn. 85. Zucha (1986: 134–35) sets up a change *-*ās* > *-*ə̄s*, followed much later by **ə̄* > **ȳ*, but both the marginal phoneme **ə̄* and the proposed sequence of *Auslautgesetze* are postulated entirely to account for the notoriously refractory ending NSl *-a* ~ SSl *-y* of the present active participle masc. nom. sg. and neut. nom./acc. sg. On more recent arguments for *-*ās* > PSl *-*ə̄*, see below in the main text.

It will be seen from the discussion here and below that I see no evidence for reflexes in Slavic of the distinction between (post-)PIE or PBSl **ā* and **ō* in final syllables, with two exceptions: nonacute (i.e., "circumflex") *-*ō* in absolute word-final position, which was raised in *n*-stem nom. sg. *-*ō* > *-*ū* > PSl *-*ȳ* (see above, fn. 6); and the diphthong *-*ōy(s)*, where the second component **y* was assimilated to **w* in PIE *o*-stem dat. sg. *-*ōy* > *-*ōw* > PSl *-*ū* (contrast *eh$_2$*-stem *-*eh$_2$-ey* > *-*āi* > PSl *-*ē*), instr. pl. *-*ōys* > *-*ōws* > *-*ū(w)s* > PSl *-*ȳ* (Meillet 1934: 153–54, 407–08, 410; Vaillant 1958: 31, 37). Note that word-final *-*s* may have already become *[h] by this stage, as in many other IE languages (Kortlandt 1979: 267).

[9] On the reconstruction of PIE *-*ōms*, see section 3 with fn. 27.

[10] For extension of acc. pl. to nom. pl. alone, cf. Schleicher 1852: 242–43, 1861–62: 433–34, 1871: 519; Miklosich 1879: 301; Brugmann 1890: 663, 675 (← *o*-stem *-*ons*), 1909: 214; Zubatý 1893: 511–12 (← *o*-stem *-*ōns*); Vondrák 1908: 8, 1928: 6; Hujer 1910: 74–76, 1920–21: 44–45; Il'inskij 1916: 359; Leskien 1919: 112; Mikkola 1950: 33–34; Rosenkranz 1955: 82; Schelesniker 1964: 37–38 (← *o*-stem *-*ons*); Kiparsky 1967: 92; Georgiev 1969: 93; Holzer 1980: 8, 10; Pohl 1983: 32, 1985: 374; Lamprecht 1987: 85, 90; Aitzetmüller 1991: 88, 90. For both nom. pl. and gen. sg., cf. J. Schmidt 1883: 338; Ljapunov 1905: 35–37; Porzeziński 1914: 108, 1916: 34, 38, 39 (← *o*-stem *-*ons*); Fortunatov 1919 [1957]: 171; Vaillant 1950: 211, 1958: 81, 83; Mareš 1964: 168 [2001: 39]; Bräuer 1969: 104, 106, 126, 127; Newman 1971: 332–33; Kortlandt 1975: 47, 1983: 180–81; Moszyński 1984: 225, 2006: 264–65; G. Schmidt 1985: 396; Zucha 1986: 136; Schenker 1993: 88, 1995: 124–25; Beekes 1995: 182–83; Townsend and Janda 1996: 144, 145, 163; Igartua 2001: 281–90, 2005a: 209–11, 227, 2005b: 290–97; Halla-Aho 2006: 159–61; Beekes 2011: 200; Janda 2014: 1567, 1569, 1572, 1577. The arguments against analogical spread of the acc. pl. ending, under the supposed influence of the feminine *i*-stems, have been ably summarized by Olander (2015: 132–33).

sim., is motivated entirely by the soft ending *-ę*. As will be argued below in section 3, the latter was not originally proper to the gen. sg. at all, so that these appeals to analogy, always improbable in any case, are likewise unnecessary.[11]

The same raising before **-s* is seen in the OCS clitic pronouns 1pl. *ny*, 2 *vy*, which have both accusative and dative value and so are more likely to continue **nōs*, **wōs* than remodeled **nŏ̄ns*, **wŏ̄ns*, with **-ns* from the nominal animate acc. pl. (cf. OPr *mans*, *wans*; see already Meillet 1897: 96).[12] Viredaz (2009: 13, 16–22) and Olander (2012: 331–32, 2015: 56, 67) independently assume a raising of pre-PSl **-ās* > **-ās* > PSl **-ə̄* > ONovg *-ě* vs. *-y* everywhere else, parallel to the development of pre-PSl **-ăs* > PSl **-ə* > ONovg *-e*, elsewhere *-ŭ*,[13] but are then forced to seek ad hoc explanations for *ny*, *vy*: either borrowing from OR (Viredaz 2009: 26) or PSl **-ə̄* > ONovg *-y* after a labial in *vy*, whence by analogy *ny* (Olander 2015: 254). The presence of the pronouns *ny*, *vy* even in the most vernacular ONovg documents rather suggests that the sequence **-ōs* > pre-PSl

[11] See Mueller 1870: 264–65; Brugmann 1890: 572–73, 1908: 191–93, 1909: 155, 158; Zubatý 1897: 23–27; Mikkola 1897: 249–50, 1950: 33; Hujer 1910: 111–12, 1920–21: 42–43; van Wijk 1916: 462; Leskien 1919: 109, 1955: 79; Noha 1924: 257; Vondrák 1924: 151–52; Rosenkranz 1955: 81; Kiparsky 1967: 83–84; Szemerényi 1970: 174, 1990: 200, 1996: 189 (**-ans* ← **-ās*); Holzer 1980: 10; Lamprecht 1987: 85–86, 90; cf. also Schleicher 1861–62: 454, 1871: 543 (gen. sg. **rankā-n(-as)* with inserted nasal, or ← acc./nom. pl.); Il'inskij 1916: 357 (gen. sg. **-āms* ← **-ās* after acc. sg. **-ām*); Otrębski 1921 (outdated comparison with other instances of secondary *-n-* in Skt and elsewhere); Sandbach 1925 (*-y* < **-ūns* < PIE **-ōn-s*; *-ę* < **-jūns* or < PIE **-ēn-s*). A version of this hypothesis takes the gen. sg. to be generalized from PIE heteroclitic *r/n*-stem gen. sg. **-en-s* → **-on-s* (Lohmann 1930; Tremblay 1996: 49, fn. 67; 50, fn. 70); but even if one derives PSl **vodā* 'water' from PIE collective **wéd-ōr*, gen. **ud-én-s*, it is hard to imagine that this and one or two other nouns of like origin could have been a sufficient basis for generalizing a new ending to all *ā*-stem nouns, adjectives, and pronouns.

Other explanations are too improbable morphologically to merit serious scrutiny: an original loc. sg. ending comparable to Ved fem. loc. sg. *tásyām* 'this' (Schelesniker 1962, 1964: 26–29, 33–34, 1967: 125, 1976; Trummer 1978: 261; Aitzetmüller 1991: 86–87, 89–90 with fn. 139; cf. already Schleicher 1852: 236; Leskien 1876: 123–24; Miklosich 1876: 4, 1879: 301); PIE gen. pl. **-om* + sg. **-s* (Knobloch 1954–55, 1956: 239–41); a combination of the two preceding (Pohl 1983: 41–42, 1985: 380–82 [gen. sg. **-ās* → loc. **-ai* + postposition **-en* + gen. sg. **-s*!]); a generalized *ū*-stem abl. sg. **-ūd/t*, formed after the model of *o*-stem **-ōd/t* (Georgiev 1969: 88–89); generalization of **-ū-* < **-ā-* to gen. sg. and nom. and acc. pl. **-ū-N* under unclear conditions (Feinberg 1978: 115); or transfer from possessive/attributive adjectives in **-n-* under Finnic influence (Greenberg 2003). Ved *-ām* in loc. sg. *tásyām*, *devyā́m*, *grīvā́yām* (to *devī́-* 'goddess', *grīvā́-* 'neck') is generally agreed to be an innovation for PInIr **-yā* (cf. Av *grīuuaiia*, OP *Aθurāyā* 'in Assyria'; Debrunner and Wackernagel 1930: 43–44 with older refs.).

[12] PBSl **nōs*, **wōs* are in turn from the PIE atonic forms **nos*, **wos* (cf. Ved *naḥ*, *vaḥ*), which stood beside tonic **n̥s-mé*, **us-wé* (Katz 1998: 96–97, 195–225; Ringe 2006: 57, 208, 2017: 70, 233). The same lengthening is found in Lat *nōs*, *vōs*.

[13] First posited by Zaliznjak (1988: 170–71).

*-*ās* became -*y* there as well, and that the ending -*ě* of the *ā*-stem gen. sg. forms is an analogical import from the soft declension (see section 2).[14]

There are thus no obstacles to assuming a development *-*ās* > *-*ūs* > PSl *-*ȳ*, and likewise *-*ą̄̆ns* > *-*ą̆̄ns* > *-*ū^{n}s* > *-*ūs* > PSl *-*ȳ* in the acc. pl. of masculine *o*-stems, with compensatory lengthening accompanying nasalization (if the vowel was first shortened in pre-PSl) followed by loss of nasality in a high vowel.[15] In contrast to the raising of *-*aN* > *-*uN*, the raising of *-*ās* > *-*ūs* must have followed the fronting of back vowels after **j* judging from the corresponding *jā*-stem forms, to which we now turn.

2. North Slavic -*ě* vs. South Slavic -*ę* in the *jo*- and *jā*-stem Declension

A complication for all treatments of Slavic *Auslautgesetze* is the ending of the *jo*-stem accusative plural and of the *jā*-stem genitive singular and nominative and accusative plural. Sobolevskij (1881) first established that, while OCS manuscripts consistently attest the ending -*ę* (e.g., acc. pl. *конѧ* 'horses'; gen. sg., nom./acc. pl. *доушѧ* 'souls', *землѧ* 'lands'), the oldest North Slavic sources have reflexes of -*ě*, namely OR -*ě* (*конѣ*, *доушѣ*), ONovg -*ě* (*koně*, *vožě* 'reins'), OCz -*ě* (*dušě* = *duſye*); cf. also Pol -*e* (*konie*, *dusze*), USor -*e* (*konje*, *duše*), Ukr -*i* (*коні*, *душі*), all continuing *-*ě*.[16] This -*ě* is sometimes referred to in Slavic historical linguistics as *ě tertium* or *$\check{e}_{3}$*, to distinguish it from the other sources of Late Common Slavic *ě (*$\check{e}_{1}$ < PBSl *ē; *$\check{e}_{2}$ < PBSl *ai*). The same divide is observable in pronominal inflection (e.g., OCS masc. acc. pl., fem. nom./acc. pl. *ję*

[14] I thank Marek Majer for clarifying this point; see his discussion in Majer 2017: 141–42, fn. 278.

[15] Forms of the 2sg. root aorist such as OCS *da*, BCS *dâ* 'gave' < PIE **deh$_3$-s*, OCS *děla* 'worked' < **eh$_2$-s* are hardly probative, since they could simply be analogical to the 3sg. (Lehr 1917: 39; Milewski 1932: 28–29; Galabov 1973: 16; Zucha 1986: 135–36; Viredaz 2009: 12; Olander 2012: 332; pace Hujer 1910: 75; Arumaa 1964: 114; Kortlandt 1979: 265, 1983: 180). On the loss of nasalization in word-final high vowels, see section 3.

Another instance of PIE *-*ōs* > PSl -*y* may be hidden in *ū*-stem nouns such as *zŭly* 'husband's sister' < **ĝl̥h$_2$ōs* (Lat *glōs* 'id.', Gk *gálōs* 'husband's sister, brother's wife'; Witczak 1998: 133–35) or *ljuby* 'love', if from an *s*-stem amphikinetic abstract **léwbh-ōs* to PIE **lewbh-* 'be pleasant' (Majer 2017: 141–45). The rest of the paradigm would then be analogical to the nom. sg., on the model of inherited PIE stems in *-*uh$_2$-* such as *svekry* 'mother-in-law', acc. *svekrŭv-ĭ*, gen. *svekrŭv-e*, etc. < PIE **swéḱruh$_2$*.

[16] ONovg -*ě* also occurs with hard stems (e.g., *o*-stem acc. pl. *kolotokě* 'wooden hammers', *ā*-stem gen. sg., acc. pl. *kuně* '*kuna*(s)', nom. pl. *kělě* 'all'); this is surely a generalization of the soft-stem ending (Vermeer 1996: 43–44, 48–51; Zaliznjak 2004: 146–47, 150; Kortlandt 2016: 91), rather than a phonologically regular development (Viredaz 2009: 16–22; Olander 2012: 333–35, 2015: 67, 132, 231, 248, 251; see above, section 1). Later Western South Slavic likewise generalizes the soft ending -*e* < *-*ę* (e.g., BCS, Sln *glave* 'heads').

'them', *moję* 'my' vs. OR *jě*, *mojě*), and thus in the definite long forms of adjectives (e.g., OCS *velikyję* vs. OR *velikyě*, ONovg *velikěě* 'great'). All four of these inflectional slots, *jo*-stem acc. pl. and *jā*-stem gen. sg., nom. pl., and acc. pl., thus show the same pattern of hard- and soft-stem endings, *-y* vs. *-ě* in North Slavic and *-y* vs. *-ę* in South Slavic.[17]

The origin of this discrepancy is a perennial problem of Slavic historical morphology, which has attracted the attention of dozens of scholars over the past 130 years.[18] The most widespread view takes both reflexes to be phonologically regular, with early denasalization of the front mid vowel in the North Slavic dialects.[19] Thus van Wijk (1916: 462–63) proposed that pre-PSl **-jens* was denasalized in WSl and ESl (**-ens* > **-ę̄s* > **-ēs* > **-ě*) but not in SSl (**-ens* > **-ę̄s* > **-ę*), while Šaxmatov (1915: 13–14) set up a Common Slavic vowel (**-öns* >) **-ě̯*, which became *i̯ę* or *i̯ä̯* and then lost its nasalization in OR *ě*, OP *e*; similarly, Fortunatov (1919 [1957]: 167–72; cf. Porzeziński 1914: 34) thought that **-jons* > **-juns* > **ų̄* > **ǭ* > **ę̄^{i}* > **i͡ę* fell together with **ę̄*, **ę* as OCS *ę*, but with **i͡e* as OR *ě*. Van Wijk's hypothesis was modified by Noha (1924: 258–63, 1927: 66–67 with fn. 1) to the effect that the change **-ens* > **-ę* took place earlier in the northern dialects and so yielded an oral vowel, but retained its nasality in the South.[20] Somewhat differently, Diels (1914) thought that **-ę* was denasalized after a "stark erweicht" consonant ("strongly softened", i.e., a yodization product), so that **-Cję* > **-C'ę* > **-C'ě*.[21] Yet another solution was suggested by Mareš (1963: 64, 1969: 110–11 [1999: 92]), for whom both **-i̯ons* and **-i̯āns* yielded an "unsystemic trimoric triphthong" **-eNZ*, which with loss of the **-Z* became SSl *-ę*, but in NSl underwent compensatory lengthening to **-ę̄* and then denasalization to *-ě*. Other authors have simply chalked up the different reflexes to phonetic fluctuation, for example Vondrák (1924: 149), who declared vaguely that "the aversion to *-ęs* appears in one part of the Slavic speech area itself to have led to *-ě* from *-ēns*, cf. OR, OBlg gen. sg. *dušě*…", or Vaillant (1950: 215, 216, 1958: 48–49, 87–88), for whom "the double treatment *-ę* and *-ě* … in origin rep-

[17] Charlie Townsend was fond of representing the correspondence NSl *ě* ~ SSl *ę* with the mnemonic symbol *ę̌*, which even has the *háček* and *ogonek* in the correct geographic positions!

[18] For useful surveys, see Arumaa 1985: 149–51; Orr 2000: 116–22; Igartua 2005a: 158–63, 199–212, 226–35; Olander 2015: 129–33, 230–31, 246–48. For the older literature, see Hujer 1910: 74–76, 98–108, 110–12, 1920–21: 47–49; Noha 1924: 244–57.

[19] See the useful survey of Orr 2000: 117–19.

[20] An early loss of nasality in East and West Slavic is also assumed by Liewehr (1955: 66–67) and Arumaa (1964: 116 **-i̯ens* > **i̯ēs* > *ě*) and given as a possibility by Bräuer (1969: 74–75, 126). Il'inskij (1916: 357–58) instead thought that PSl **-'ens* lost the nasal early in WSl and ESl, followed by compensatory lengthening of **-es* > *-ě*.

[21] Quotes in this paragraph, in fn 22, and in fn 34 have been translated from the original into English by the author.

resented only a rather slight variation, because *-ě* was the denasalized form of *-ę*: surely a simple vacillation of pronunciation which only afterwards became a dialectal divergence" (Vaillant 1958: 48–49; followed by Kiparsky 1967: 56, 84).[22]

Aside from their evidently ad hoc nature, all of these treatments require various analogical explanations for the pan-Slavic masc./neut. nom. sg. *-ę* in the present active participle of *je-* and *i*-presents, which can only continue a final sequence of the shape pre-PBSl **-jants*, **-ints*.[23] Furthermore, they do not address the question why there was no parallel tendency to denasalize the nom./acc. sg. ending *-ę* of the neuter *n-* and *nt*-stem nouns, or for that matter the back nasal vowel **ǫ* in word-final position. As decisively argued by Nitsch (1928), there is also no way to derive OPol *-e* from PSl **-ę*, which otherwise always retains its nasality, even in the accusative clitics 1sg. *mię*, 2sg. *cię*, 3sg. *się*.

More recently, Viredaz (2009: 22–23) reconstructs a separate vowel for Proto-Slavic: pre-PSl **-ās*, **-ans* > PSl **-ā* > ONovg *-ě*, elsewhere *-y*; the allophone after soft consonants became NSl *-ě*, but SSl *-ę* with "spontaneous nasalization", which is hardly less ad hoc than the assumption of early denasalization in NSl. Olander (2015: 56), apparently independently of Viredaz, proposes a similar development of **-āns*, **-ōns* > PSl **-ān* > ONovg *-ě*, elsewhere *-y*, but thinks that the soft counterpart gave both dialectal variants directly: **-jāns*, **-jōns* > PSl **-i̯ān* > SSl *-ję* ~ NSl *-jě* (Olander 2012: 334, 2015: 248, 251). The *jā*-stem acc. pl. was extended to the nom. pl. and gen. sg. in South Slavic, replacing the expected reflex of **-jās* > PSl **-i̯ā* (> NSl *-jě*). However, the ONovg hard-declension forms in *-ě* are more simply explained as analogical transfers from the *jo-* and *jā*-stem declension (see above, fn. 16), and the evidence of ONovg *ny*, *vy* also argues for a unitary outcome of (post-)PIE **-ā(n)s*, **-ō(n)s* > **-ā(n)s* > **-ū(n)s* > PSl **-ȳ* (section 1). In any case, the assumed dialectal distribution of the reflexes *-ję* and *-jě* is ad hoc: nowhere else in Olander's model of Slavic *Auslautgesetze* is there a split between northern and southern treatments of a word-final sequence.

Since it is not possible to derive both soft-stem endings from a single preform, other scholars have proposed morphological explanations for the co-

[22] Similarly Zucha 1986: 134 (under "Denasalierung *i̯*, *u̯* → *ī*, *ū*": "In SSl **daušjęs* was preserved, while in NSl this too was denasalized to **daušjēs*"), Townsend and Janda 1996: 52–53 ("[i]n SSl…the result is *ę*, regardless of position, but in WSl and ESl (= NSl), this vowel could denasalize to *ě* (also known as *ě*$_3$) in final position").

[23] This holds no matter how one explains the puzzling split of NSl *-a* vs. SSl *-y* in the present active participle of *e*-presents, another notorious crux of Slavic historical grammar. See the useful survey by Olander (2015: 88–93), although I cannot follow his own proposal of a development PIE **-(y)onts* > PSl **-(j)ən*.

existence of *-ě* and *-ę* in the oldest Slavic languages.[24] One subset of these, going back to Jagić (1893: 522–23), sees **-ę* as the original PSl ending in all the inflectional slots, namely the gen. sg., nom. pl., and acc. pl. of *jā*-stems as well as the acc. pl. of *jo*-stems. The reflexes of **-ę* were replaced with *-ě* in the NSl languages to avoid syncretism with the *jā*-stem nom. sg. and *jo*-stem gen. and animate acc. sg. in **-a*. But this approach, adopted among others by Furdal (1961: 60–62) and Newman (1971: 333–37), is based on several unverifiable assumptions about prehistoric Slavic phonetics, namely, that **ę* was a low front [ä] (i.e., [æ̃]) and so following denasalization was in danger of falling together with **a* [ä] after a palatalized consonant, or that **-ę̄* merged in the Lechitic dialects with the acc. sg. and inst. sg. to create an undesirable homophony in nouns of the *wola*-type (thus gen. sg., nom./acc. pl. **wol'ę̄*> **wol'ǭ* like acc. sg. **wol'ǭ*, inst. sg. **wol'ejǭ* > **wol'ǭ*; Newman 1971: 334–37). Even supposing that speakers of OR, OCz, or OP felt the need to eliminate the case-number syncretism in these nouns, why did they introduce *-ě*, rather than simply take over *-y* from the hard *ā*-stem declension?[25]

Alternative analogical accounts are no less problematic. Schelesniker (1962, 1964: 26–29, 1967: 125, 1976; cf. Igartua 2005a: 211–12) took NSl *-ě* to be the regular sound-change outcome of pre-PSl **-jās*, but SSl *-ę* (and pan-Slavic *-y*) from an alleged locative ending **-ōm* comparable to Skt fem. loc. sg. *tasyām*; aside from the functional difficulties and the unfounded assumption that Slavic did not inherit the genitive from PIE, the latter ending is almost certainly an Indo-Aryan innovation (see above, fn. 11 ad fin.). Shevelov (1964: 334–35) suggested that NSl *jo*-stem acc. pl. *-ě* goes back to an apocopated **-ē* ← **-ēN*, with final nasal dropped after the other *jo*-stem case forms; but this ad hoc analogy hinges on the author's idiosyncratic view of word-final nasals and in any case will not work for the *ā*-stems, where both the acc. sg. and inst. sg. retained their nasal vowel into PSl.[26] Schmalstieg (1968: 48–50; 1971:

[24] The suggestion that SSl *-ę* vs. NSl *-ě* goes back to fluctuation between **-ns* and **-s* in PSl or even PIE (see e.g., Endzelīns 1913: 112, Otrębski 1921: 12–13) is rooted in outdated conceptions of IE historical morphology. On the reconstruction of the PIE animate acc. pl. endings, see section 3 below.

[25] I do not share the conventional view that **-jā(-)* was fronted to and fell together with **-jē(-),* followed by analogical restoration to **-ā(-)* in *jā*-stem nom. sg. **-ā*, acc. sg. **-ān* > **-ą̄*, dat./inst. du. **-āmā*, dat. pl. **-āmu*, inst. pl. **-āmī*, loc. pl. **-āsu* (OCS *-a*, *-ǫ*, *-ama*, *-amŭ*, *-ami*, *-axŭ*) as well as *jo*-stem gen. sg. **-ā* and masc. nom./acc. du. **-ā* (OCS *-a*). Spellings with *ě* in Glagolitic OCS manuscripts surely indicate a **phonetic** fronting of PSl **ā*, but not **phonemic** merger with the usual reflex of PSl **ē* (i.e., *ě*). As for the backing of **ě* > **a* after **č*, **š*, **ž* in for example OCS, OR *ležati* 'lie', *časŭ* 'time', this sound change could have been restricted to nonfinal position, hence its absence in OR, OCz gen. sg., nom./acc. pl. *dušě*.

[26] Interestingly, he speculates that the failure of this analogy in SSl "might have been due to the earlier rise of nasal vowels in the South" (Shevelov 1964: 335), exactly the

142–43) identified the source of NSl *-ě* vs. SSl *-ę* in the long-form adjective, for example **sin'-ens-ens* 'blue' > **sin'-ens-ēs* > **sin'-ęs-ē* → NSl **-ějě* (cf. **zemj-ě* 'earth') but SSl **-ęję* (whence **zemj-ē* → **zemj-ę*), but this explanation requires a long series of often improbable analogical changes; furthermore, soft-stem adjectives were proportionally much rarer than *jo*- and *jā*-stem nouns, making the presumed analogy even less likely. Georgiev (1969: 98–99) similarly thought that NSl *-ě* and SSl *-ę* could continue sandhi variants **-jās* and **-jāns* of the acc. pl. ending, which was extended to the nom. pl. and then to the gen. sg. (for invariant **-jās*); but the assumed sandhi development **-Vns C-* > **-Vs C-* is unsupported by independent examples, and the survival of isofunctional allomorphs for any length of time is inherently improbable and should be considered only as a last resort.

3. Forward Reconstruction and a New Proposal

The only remaining possibility—one that has rarely been considered in the literature to date—is that one or more of the inflectional slots in question originally had **-ě* (or its predecessor), the other(s) **-ę* (or its predecessor), and different Slavic dialect areas generalized different endings, producing a uniform alternation *-y ~ -ě* in the North and *-y ~ -ę* in the South. The Slavic languages themselves offer no hints of a distinction between **-ě* and **-ę* in any of the relevant case forms, but the Proto-Slavic situation can be recovered with some confidence if we reconstruct forward from PIE. The PIE accusative plural of animate *o*-stems is usually given as **-oms*, but certain forms rather point to a preform **-ōms* (e.g., Ved *-ān* and its sandhi variants *-āṃ V-*, *-āṃś c-*, *-āṃs t-*, or Lith *-us*, *-úos-ius*).[27] For eh_2-stems, the genitive singular goes back to PIE **-eh_2-es* (> Lith *-ōs*, Gk *-ᾶ̃s/-εᾶ̃s*), as does the nominative plural (> Lith *-os*). The accusative plural involves a complication, however: the expected shape of the ending is ***-eh_2-ms*, but it has long been assumed that this sequence was reduced already in PIE to **-ās* (J. Schmidt 1883: 337–38; Brugmann 1909: 225–26), or in modern laryngeal notation **-eh_2s*.[28] The latter is directly continued in Indo-Iranian (Ved *-āḥ*, Av *-å*), Germanic (Goth *-ōs*), and Italic (South Picene

opposite of the view held by Noha and others (see above with fn. 20).

[27] See Kim 2012; Olander 2015: 250; otherwise Kortlandt 2016: 92–93 (PIE **-ons* with laryngeal adopted from **-aHns* in BSl). A long-vowel ending was proposed for PIE already by Hanssen (1885: 615, fn. 1 [Gk *-ons* < **-ōns*]; cf. Zubatý 1893: 508, 1896: 273, fn. 1; Vondrák 1898: 337–38, 1904: 190–91; Hujer 1910: 106), but was rejected by Brugmann (1890: 672, 1909: 224–25) and thereafter seems to have passed out of currency except among Balticists (cf. Stang 1965: 294–95, 1966: 186; Zinkevičius 1980 I: 211 [**-ōns*]). On the Baltic forms, see section 4.

[28] On the consequences for pre-PIE phonology of the deletion of ***m* in ***-eh_2ms* > **-eh_2s*, see Kim forthcoming b.

-as), but was remodeled after other nonneuter stems in Greek (*-*ās* → *-*āns* > Proto-Greek *-*ans* > Cretan -*ans*, Attic-Ionic -*ās*, Lesbian -*ais*, etc.) and, following the Proto-Sabellic change of word-final *-*ns* > *-*f*, in Oscan (*-*ās* → *-*āfs* > **-ass**) and Umbrian (*-*ās* → *-*āf* > **-af**).[29]

It is likely that Slavic, like Greek and Sabellic, also remodeled the eh_2-stem accusative plural ending to *-*āNs* for the following reason. If PIE acc. pl. *-eh_2*s* was inherited unchanged in Balto-Slavic, it would have become *-*ā́s*, contrasting only in intonation with nonacute ("circumflex") *-*ās* in the gen. sg. and nom. pl. Since there is no evidence that intonation would have affected the evolution of word-final sequences *-*V̄s* in Slavic, all three of these case endings should have developed by sound change to -*y* in hard stems (see above) and -*ě* in soft stems (Table 4).

Table 4

	PIE	PBSl	PSl
jā-stem gen. sg.	*-yeh₂es	> *-jās	> *-jēs > *-ē
jā-stem nom. pl.	*-yeh₂es	> *-jās	> *-jēs > *-ē
jā-stem acc. pl.	*-yeh₂s	> *-já́s	> *-jēs > *-ē

But in that case, the only source for the SSl ending -*ę* in the soft *jā*-stem declension would have been the accusative plural of *jo*-stems, where *-*ę̄* is the expected outcome of *-*jens* < pre-PSl *-*jans*. One would thus have to assume that SSl introduced a new allomorph -*ę* into the acc. pl. of *jā*-stems by interparadigmatic analogy to the *jo*-stems, then extended this ending to the gen. sg. and nom. pl., whereas NSl conversely replaced the acc. pl. of *jo*-stems with that of the *jā*-stems (Zubatý 1893: 511–15; cf. Vondrák 1904: 189–91, 1906: 53, 108, 1908: 3; Hujer 1910: 99–105, 1920–21: 37; Fortunatov 1919 [1957]: 169–72; Schelesniker 1964: 35–40).[30]

Although this scenario cannot be ruled out, the variation between NSl -*ě* and SSl -*ę* finds a much more straightforward diachronic motivation if there was a contrast of endings **within** the *jā*-stems already in Proto-Slavic. Since

[29] On the Italic endings, see Rix 1986. The ambiguous Latin ending -*ās* therefore probably also continues Proto-Italic *-*ās* (Meiser 1998: 133; Weiss 2011: 236).

[30] This was apparently also the opinion of Aitzetmüller (1991: 84, 90–91) for the SSl *jā*-stem nom. and acc. pl. -*ę*, but he connected gen. sg. -*ę* with the Skt loc. sg. -*ām*, which is an innovation of Indo-Aryan (see above, fn. 11 *ad fin.*). Zubatý (1893: 516–17) suggested that NSl replaced *jo*-stem acc. pl. *-*ę* with -*ě* to avoid homophony with nom. sg. *-*(j)a*, whereas SSl replaced *jā*-stem acc. pl. *-*ě* with -*ę* (whence also nom. pl., gen. sg. -*ę*) for the same reason. On Jagić's modification of this hypothesis and its weaknesses, see section 2.

there can be no question of a phonological source for SSl *-ę* in the gen. sg. or nom. pl., the original distribution must have been **-ē* in those cases and **-ę̄* in the accusative plural. The latter would then continue **-jēns* < **-jāns* ← **-jās* < PIE **-yeh$_2$s*, with the same kind of remodeling as in Proto-Greek **-ans* or the Sabellic languages (Meillet 1897: 125–26, 1914–15: 6–7, 1934: 398–99; Vaillant 1958: 83–84). Crucially, the hard-stem endings all merged as **-ȳ*, with loss of nasality in high vowels as in the acc. pl. of *i*- and *u*-stems: PIE **-ims*, **-ums* > PBSl **-ins*, **-uns* > **-īs*, **-ūs* > PSl **-ī*, **-ȳ* (OCS *-i*, *-y*).[31] The evolution of the three case forms of *ā*-stems, along with the acc. pl. of *o*-stems, was thus as in Table 5.

Table 5

	PIE	PBSl			PSl
ā-stem gen. sg.	**-eh$_2$es* >	**-ās* >		**-ūs* >	**-ȳ*
ā-stem nom. pl.	**-eh$_2$es* >	**-ās* >		**-ūs* >	**-ȳ*
ā-stem acc. pl.	**-eh$_2$s* →	**-ā́ns* >		**-ūns* > **-ūs* >	**-ȳ*
o-stem acc. pl.	**-ōms* >	**-ōns* > **-āns* >		**-ūns* > **-ūs* >	**-ȳ*
jā-stem gen. sg.	**-yeh$_2$es* >	**-jās* >	**-jēs* >		**-ē*
jā-stem nom. pl.	**-yeh$_2$es* >	**-jās* >	**-jēs* >		**-ē*
jā-stem acc. pl.	**-yeh$_2$s* →	**-já́ns* >	**-jēns* >		**-ę̄*
jo-stem acc. pl.	**-yōms* >	**-jōns* > **-jāns* >	**-jēns* >		**-ę̄*

As this table shows, I assume that raising of **ă* > **ŭ* before nasals preceded fronting of back vowels after **j*, which in turn preceded raising of final **-ās* > **-ūs* (section 1 ad fin.). Whatever the phonetic realization of the pre-PSl acc. pl. ending **-āns*, the sequence of vowel + nasal was treated as phonemically equivalent to **ā* and so escaped the first raising, only later becoming **-ūns*. On the other hand, the soft endings underwent fronting but not the second raising: **-jās* > **-jēs* > **-ē*, **-jāns* > **-jēns* > **-ę̄*.[32] Following the dissolution of Proto-

[31] See Halla-Aho 2006: 144–59 and Olander 2015: 243–46, with a review of competing views. Lunt (2001: 227, 229) assumes a second round of denasalization in **-ŏns*, **-āns* > **-ūns* > **-ūns* ~ **-jīns* > **-ū* ~ **-jīn* > OCS *-y* ~ *-ę*, but there is no reason to separate the hard *o*- and *ā*-stem acc. pl. ending from that of the *u*-stem acc. pl., and the lowering of the high nasal diphthong **īN* > **ēN* (and loss of nasality in NSl) here but not in the *i*-stem acc. pl. is *ad hoc*.

[32] Pace Matasović (2008: 124–25, 126, 146), raising followed by fronting would have given **-jāns* > **-jūns* > **-jīns* > PSl **-C′ī*, with the same denasalization as in *i*-stem acc. pl. **-i*. For alternative relative chronologies of the relevant sound changes, see among many others Shevelov 1964; Holzer 1980; Kortlandt 1983, 1994; Zucha 1986: 134–35; Lunt 2001: 196, 207–08; and Olander 2015: 46–67.

Slavic, the northern and southern dialects generalized respectively *-*ē* (> OR, OCz -*ě*) and *-*ę̄* (> OCS -*ę*).

Although scholars have long assumed generalization of the *jā*-stem acc. pl. reflex to the nom. pl. and gen. sg. in OCS -*ę*,[33] the only authors I have found who explicitly state that the NSl dialects must have leveled -*ě* in the opposite direction are Galabov (1973: 10–11) and Moszyński (1984: 225, 286–87, 2006: 264–65, 339–40), neither of whom refers to the wider IE context.[34] The crucial point is that the divergence between the northern and southern reflexes can be motivated far more easily by reconstructing two separate endings *within jā*-stem inflection for Proto-Slavic: *-*ē* in the gen. sg. and nom. pl., and *-*ę̄* in the acc. pl. The elimination of this contrast in all Slavic dialects is a typical example of grammar simplification through reduction of the number of morphophonologically conditioned alternations: in place of gen. sg./nom. pl. *-*ȳ* ~ *-*ē* and acc. pl. *-*ȳ* ~ *-*ę̄* in the hard and soft declensions, respectively, OCS and the rest of South Slavic generalized the unitary alternation -*y* ~ -*ę*, while OR, OCz, and other North Slavic dialects instead opted for -*y* ~ -*ě*.[35] The reconstruction of PSl *-*ę̄* for the *jā*-stem acc. pl. in turn implies that the PIE ending *-*ās* was remodeled to *-*āns* in the prehistory of Slavic, much as in Greek and the Sabellic languages.

4. A Balto-Slavic Date for the Remodeling?

It is possible that the remodeling of the *ā*-stem accusative plural ending to *-*āns* took place already in Proto-Balto-Slavic, but the Baltic facts are complex and difficult to interpret. Standard Lith -*às* could be from either *-*ą́s* or *-*ą́ns* with shortening by Leskien's Law, but since -*ąs*- in the definite adjective ending -*ą́s-ias* can only continue *-*ą́ns*, it is natural to assume the same immediate

[33] For example, Meillet 1934: 398; Tedesco 1951: 173; Mathiassen 1989: 124; Rasmussen 1992 [1999]: 507, fn. 2; Olander 2015: 131–32, 231. See also the references above in fn. 10.

[34] Cf. Igartua 2005a: 234–35 (tentative), citing Galabov; and Schenker 1995: 125, citing Moszyński. Milewski (1932: 7–8) came close in proposing *-*āns* > *-*ās* > -*y* || -*ʼě* in line with a general "phonetic tendency" of PSl, which in the SSl dialects was "sparaliżowany wpływem analogji morfologicznej" ("paralyzed by the influence of morphological analogy", presumably of the *o*-stem acc. pl. in *-*ons*) to give *-*ans* > -*y* || -*ʼę*; as a result, SSl generalized -*ę* from the *jo*- and *jā*-stem acc. pl. to the *jā*-stem gen. sg. and nom. pl., while NSl instead extended -*ě* from the *jā*-stem gen. sg. and nom./acc. pl. to the *jo*-stem acc. pl. In contrast to Milewski, I assume remodeling of the *ā*-stem acc. pl. ending for PSl, if not already PBSl (see below, section 4).

[35] Cf. Ferrell 1971: 90–93, although I cannot follow his explanations for the spread of -*ě* in the North and -*ę* in the South. In contrast, the isolated alternation *-*ȳ* ~ *-*ī* (< PBSl *-*ōis* < post-PIE *-*ōys*; see fn. 8) in the *o*-stem inst. pl. remained unaffected, until it was lost by sound change and/or analogical remodeling in all Slavic dialects.

source for *-às*, rather than positing coexistence of both variants **-ā́s* and **-ā́ns* in pre-Lithuanian (Fortunatov 1919 [1957]: 170–71; Endzelīns 1948 [1971]: 145, revising 1923: 308). OPr *-ans* in *gennans* 'women', *rānkans* 'hands' also points to a PBSl date for the remodeling.

On the other hand, acc. pl. *-as*, *-os-ias* are found in eastern Lithuanian dialects and much of the Žemaitian dialect area (Zinkevičius 1980 I: 194). These endings, along with loc. pl. *-os-e* and the archaic illative pl. *-os-na*, require a preform **-ās*, as does Latv *-as*, definite *-ās* (J. Schmidt 1883: 337–38; Brugmann 1890: 674, 675, 1909: 224–25, 226; Hujer 1910: 99; Stang 1966: 200), which would imply that the nasalless ending reconstructed for PIE survived unchanged into PBSl after all, and the remodeling to **-āns* took place independently in Slavic, West Baltic, and parts of Lithuanian;[36] or alternatively, that there was variation in PBSl between **-āns* and **-ās*, which was resolved one way or the other in the daughter languages (Vaillant 1958: 83; Georgiev 1969: 45). However, these forms too may be explained as innovations, for example by analogy to the short ending *-as* on the model of nom. sg. *-a* ~ *-o-* (Mathiassen 1989: 124); or in the case of Latvian, by analogy to the *i*-stems, where the nom. and acc. pl. endings are identical (Rasmussen 1992 [1999]: 507, fn. 2).

Most recently, Olander (2015: 248) also reconstructs PBSl **-āns*, with "the Baltic forms pointing to a proto-form without a nasal hav[ing] arisen at a relatively late stage, as the result of an East Baltic loss of **n* between a long vowel and a final **s*".[37] The loss would have created allomorphy between the substantives and indefinite (short-form) adjectives on the one hand and the definite (long-form) adjectives on the other, which was retained in standard Lith *-ąs-ias*, but leveled in dialectal *-os-ias* and Latv *-ãs*. As for Lith loc. pl. *-os-e* and ill. pl. *-os-na*, they could well have been created after the Proto-East-Baltic stage, with the vowel subsequently remaining unaffected by Leskien's Law. This explanation will also work for *o*-stem acc. pl. PBSl **-ōns* > **-ōs*: the suffixed allomorph is reflected in Žemaitian *-úns-ius*, *-ū́s-ius* and loc. pl. *-uns-è*, *-ūs-è* (Zinkevičius 1980 I: 211), while the standard form *-úos(-ius)* has lost the nasal in absolute final position, there being no nasal diphthong **[ũõ]* (Mathiassen 1989: 124–25). These developments may be summarized in Table 6 on the following page; analogical forms are in brackets.

[36] See for example Weiss 2011: 236 and now Jasanoff 2017: 139, fn. 29: "It is minimally clear that some dialects never had **-n-* in the *ā*-stems."

[37] Olander (2009: 183, 2015: 242) sets up a phonetic rule whereby word-final **-ns* caused glottalization of a preceding vowel; see already Streitberg (1894), who proposed a rule lengthening vowels before **-ns* to account for the *o*- and *ā*-stem acc. pl. endings. However, the acute intonation of Lith *-ùs*, *-úos-* and *-às*, *-ą́s-* can directly continue the long vowel of the PBSl preforms **-ōns*, **-āns* (Kim 2012: 150–51).

Table 6

	PBSl		Proto-East-Baltic		
ā-stem acc. pl.	**-ā́ns*	>	**-ā́s #*	>	Lith *-as*, loc. *-os-e*, ill. *-os-na* [dial. *-os(-ias)*]
					Latv *-as* [*-ãs*]
			**-ā́ns-*	>	Lith *-ą́s(-ias)*
o-stem acc. pl.	**-ṓns*	>	**-ṓs #*	>	Lith *-úos(-ius)*, Latv *-us*, *-uõs*
	**-ṓns-*	>	**-ṓns-*	>	Lith dial. *-uns(-ius)*, *-ū́s(-ius)*, loc. *-uns-è*, *-ūs-è*

5. Summary

The correspondence NSl *-ě* ~ SSl *-ę* in the *jo*-stem accusative plural and *jā*-stem genitive singular and nominative and accusative plural is best explained by positing a PSl contrast within the soft *jā*-stems between gen. sg., nom. pl. **-ē* and acc. pl. **-ę̄*, which was leveled in different directions as NSl *-ě* and SSl *-ę*. With its nasal vowel, the acc. pl. ending must go back to **-jāns*, thereby demonstrating that the PIE eh_2-stem desinence **-eh_2s* > **-ās* was remade to **-āns* after the other declensional classes in the early prehistory of Slavic, as in Greek and Sabellic. The Baltic endings, which on the surface continue both **-ā́s* and **-ā́ns*, are consistent with a PBSl date for this innovation.

The Slavic *jā*-stem accusative plural thus presents another example of an all too familiar problem in morphological reconstruction: the data from the languages belonging to a particular branch require the reconstruction of a morphological contrast (or alternation) for their common ancestor, here PSl, but the exact distribution of that contrast is simply unrecoverable without reference to data from related branches, in this case of the larger Indo-European family. The solution proposed here raises the possibility that other old cruxes of Slavic historical morphology may similarly find an explanation within an Indo-European context, for example the divergence in the masc./neut. nom. sg. of *e*-present active participles between NSl *-a* and SSl *-y*, or vowel alternations such as OCS, Rus *četyre* vs. Pol *cztery*, Cz *čtyři* 'four' or OCS *tysǫšti*, Cr *tisuća* vs. OCS *tysęšti*, Rus *tysjača*, Pol *tysiąc* 'thousand'.

References

Aitzetmüller, Rudolf. (1991) *Altbulgarische Grammatik als Einführung in die slavische Sprachwissenschaft*. 2., verbesserte und erweiterte Auflage. Freiburg i. Br.: Weiher. [*Monumenta Linguae Slavicae Dialecti Veteris. Fontes et Dissertationes*, Tom. XXX.]

Arumaa, Peeter. (1964) *Urslavische Grammatik. Einführung in das vergleichende Studium der slavischen Sprachen*. I. Band: *Einleitung. Lautlehre*. Heidelberg: Winter.

————. (1985) *Urslavische Grammatik. Einführung in das vergleichende Studium der slavischen Sprachen*. III. Band: *Formenlehre*. Heidelberg: Winter.

Beekes, Robert S. P. (1995) *Comparative Indo-European linguistics: An introduction*. Amsterdam: John Benjamins.

————. (2011) *Comparative Indo-European linguistics: An introduction*. 2nd ed., revised and corrected by M. de Vaan. Amsterdam: John Benjamins.

Birnbaum, Henrik. (1998) "Common Slavic in time and space". *Scando-Slavica* 44(1): 131–43.

Bräuer, Herbert. (1969) *Slavische Sprachwissenschaft*. II: *Formenlehre*. 1. Teil. Berlin: Walter de Gruyter.

Brugmann, Karl. (1890) *Grundriss der vergleichenden Grammatik der indogermanischen Sprachen*. Zweiter Band: *Wortbildungslehre (Stammbildungs- und Flexionslehre)*. Zweite Hälfte, Erste Lieferung: *Zahlwortbildung, Casusbildung der Nomina (Nominaldeclination), Pronomina*. Strassburg: Trübner.

————. (1908) "Die Anomalien in der Flexion yon griech. γυνή, armen. *kin* und altnord. *kona*". *Indogermanische Forschungen* 22(1): 171–93.

————. (1909) *Grundriss der vergleichenden Grammatik der indogermanischen Sprachen*. Zweite Bearbeitung. Zweiter Band: *Lehre von den Wortformen und ihrem Gebrauch*. Zweiter Teil. 1. Lieferung: *Zahlwörter. Die drei Nominalgenera. Kasus- und Numerusbildung der Nomina. Pronominalstämme und Kasus- und Numerusbildung der Pronomina*. Strassburg: Trübner.

Debrunner, Albert and Jakob Wackernagel. (1930) *Altindische Grammatik*. Band III: *Nominalflexion – Zahlwort – Pronomen*. Göttingen: Vandenhoeck & Ruprecht.

Diels, Paul. (1914) "*-ě* und *-ę* in den Endungen der slavischen Deklination". *Archiv für slavische Philologie* 35: 321–24.

Endzelīns, Jānis. (1913) "Lingvističeskīja meloči. 1: O proisxoždenīi drevne-russkago i zapadno-slavjanskago -ě, sootvětstvujuščago južno-slavjanskomu -ę. 2: Bolg. pendeset i devendeset, i nd iz" dd". *Russkīj filologičeskīj věstnik"* 70(1): 109–13.

————. (1923) *Lettische Grammatik*. Heidelberg: Winter.

————. (1948) *Baltu valodu skaņas un formas*. Rīga: Latvijas valsts izdevniecība.

————. (1971) *Comparative phonology and morphology of the Baltic languages*. Tr. by William R. Schmalstieg and Benjamiņš Jēgers. The Hague and Paris: Mouton. [*Slavistic printings and reprintings*, 85.]

Feinberg, Lawrence E. (1978) "Thematic vowel alternation in Common Slavic declension". Ernest A. Scatton, ed. *Studies in honor of Horace G. Lunt on the occasion of his 60 birthday, Sept. 12, 1978*. Columbus, OH: Slavica Publishers, 107–21. [= *Folia slavica* 2(1–3).]

Ferrell, James. (1971) "On the Slavic nom. sg. masculine and neuter of the present active participle and the problem of *ě* tertium". Bohuslav Havránek, ed. *Studia palaeoslovenica: Josepho Kurz septuagenario dedicatum*. Prague: Academia, 85–93.

Fortunatov, F. F. (1919) *Lekcīi po fonetikě staroslavjanskago (cerkovnoslavjanskago) jazyka*. Petrograd: Tip[ografija] Rossījskoj Akademīi Nauk. [Reprinted in Fortunatov 1957, 5–256.]

————. (1957) *Izbrannye trudy*. Tom 2. Moscow: Gosudarstvennoe učebno-pedagogičeskoe izdatel'stvo Ministerstva Prosveščenija RSFSR.

Furdal, Antoni. (1961) *Rozpad języka prasłowiańskiego w świetle rozwoju głosowego*. Wrocław: Zakład Narodowy im. Ossolińskich. [*Prace Wrocławskiego Towarzystwa Naukowego, Seria A*, 70.]

Galabov, Ivan. (1973) "Urslavische Auslautprobleme". *Wiener slavistisches Jahrbuch* 18: 5–17.

Georgiev, Vladimir. (1969) *Osnovni problemi na slavjanskata diaxronna morfologija*. Sofia: Bălgarska Akademija na Naukite, Institut za Bălgarski Ezik.

Greenberg, Marc L. (2003) "On the possible Uralic source for the Slavic gen. sg. ā-stem desinence in Slavic". Ju. V. Anduganov et al., eds. *Aktual'nye problemy finno-ugorskoj filologii*: *Materialy meždunarodnoj naučnoj konferencii, posvjaščennoj 70-letiju professora I. S. Galkina*. Joškar-Ola: Marijskij Gosudarstvennyj Universitet, 44–50.

Halla-Aho, Jussi. (2006) *Problems of Proto-Slavic historical nominal morphology on the basis of Old Church Slavic*. Ph.D. dissertation, University of Helsinki.

Hanssen, Friedrich. (1885) "Der griechische circumflex stammt aus der ursprache". *Zeitschrift für vergleichende Sprachforschung* 27(6): 612–17.

Hirt, Herman. (1893) "Zu den slavischen Auslautsgesetzen". *Indogermanische Forschungen* 2: 337–64.

Holzer, Georg. (1980) "Die urslavischen Auslautgesetze". *Wiener slavistisches Jahrbuch* 26: 7–27.

————. (1998) "Zur Rekonstruktion urslavischer Lautungen". Jerzy Rusek and Wiesław Boryś, eds. *Prasłowiańszczyzna i jej rozpad*. Warsaw: Energeia, 57–72.

————. (2003) "Urslavische Phonologie". *Wiener slavistisches Jahrbuch* 49: 23–40.

Hujer, Oldřich. (1910) *Slovanská deklinace jmenná*. Prague: Česká Akademie Císaře Františka Josefa pro vědy, slovesnost a umění.

————. (1920–21) *Praslovanské tvarosloví*. Část I: *Praslovanské skloňování (výtah z univ. čtení)*. Prague: Spolka Československých Filosofů a Přírodovědců.

Igartua Ugarte, Iván. (2001) "El origen del genitivo singular de los temas en *-ā* en eslavo. Una incursión en la morfología interna del eslavo común". *Eslavística complutense* 1: 269–300.

Igartua Ugarte, Iván. (2005a) *Origen y evolución de la flexión nominal eslava*. Bilbao: Euskal Herriko Unibertsitatea / Universidad del País Vasco. [*Filología y lingüística*, 11.]

————. (2005b) "Structural analogy and the inflectional fate of the Slavic **ā*-stems". *Indogermanische Forschungen* 110: 282–307.

Il'inskij, G. A. (1916) *Praslavjanskaja grammatika*. Nežin": Pečatnik".

Jagić, Vatroslav. (1893) "Ein Zusatz" [response to Zubatý 1893]. *Archiv für slavische Philologie* 15: 518–24.

Janda, Laura A. (2014) "Introduction to Slavic historical morphology: Slavic noun classes". Timan Berger, Karl Gutschmidt, Sebasian Kempgen, and Peter Kosta, eds. *Die slavischen Sprachen. Ein internationales Handbuch zu ihrer Struktur, ihrer Geschichte und ihrer Erforschung / The Slavic languages. An international handbook of their history, their structure, and their investigation*. Vol. 2. Berlin and New York: de Gruyter Mouton, 1565–82. [*Handbücher zur Sprach- und Kommunikationswissenschaft*, 33/2.]

Jasanoff, Jay H. (1983) "A rule of final syllables in Slavic". *Journal of Indo-European studies* 11: 139–49.

————. (2014) "A note on the Slavic genitive plural". Michael S. Flier, David J. Birnbaum, and Cynthia M. Vakareliyska, eds. *Philology broad and deep: In memoriam Horace G. Lunt*. Bloomington, IN: Slavica Publishers, 143–50.

————. (2017) *The prehistory of the Balto-Slavic accent*. Leiden and Boston: Brill. [*Brill's Studies in Indo-European Languages & Linguistics*, 17.]

Katz, Joshua T. (1998) *Topics in Indo-European personal pronouns*. Ph.D. dissertation, Harvard University.

Kim, Ronald I. (2012) "The PIE thematic animate accusative plural revisited". Roman Sukač and Ondřej Šefčík, eds. *The sound of Indo-European 2: Papers on Indo-European phonetics, phonemics and morphophonemics*. Munich: LINCOM, 144–58. [*LINCOM Studies in Indo-European Linguistics*, 41.]

————. (Forthcoming a) "Problems in the reconstruction of Proto-Slavonic phonology". Jan Fellerer and Neil Bermel, eds. *The Oxford guide to the Slavonic languages*. Oxford: Oxford University Press.

————. (Forthcoming b) "Stang's Law and the phonological status of **m* in Proto-Indo-European".

Kiparsky, Valentin. (1967) *Russische historische Grammatik*. Band II: *Die Entwicklung des Formensystems*. Heidelberg: Winter.

Knobloch, Johann. (1954–55) "Zur Erklärung des Genitivs Sing. fem. auf *-y* im Slawischen". *Wissenschaftliche Zeitschrift der Ernst-Moritz-Arndt-Universität Greifswald, Gesellschafts- und sprachwissenschaftliche Reihe* 4: 255–56.

————. (1956) "Der Ausdruck des partitiven Verhältnisses im Slawischen und in einigen verwandten Sprachen". Hans Holm Bielfeldt, ed. *Vorträge auf der Berliner Slawistentagung (11.–13. November 1954)*. Berlin: Akademie-Verlag, 233–42. [*Veröffentlichungen des Instituts für Slawistik*, 8.]

Kortlandt, Frederik. (1975) *Slavic accentuation: A study in relative chronology*. Lisse: de Ridder.

————. (1978) "On the history of the genitive plural in Slavic, Baltic, Germanic, and Indo-European". *Lingua* 45: 281–300. [Reprinted in Kortlandt 2009, 111–23.]

————. (1979) "On the history of the Slavic nasal vowels". *Indogermanische Forschungen* 84: 259–72. [Reprinted in Kortlandt 2011, 99–109.]

————. (1983) "On final syllables in Slavic". *Journal of Indo-European studies* 11: 167–85. [Reprinted in Kortlandt 2011, 119–33.]

————. (1994) "From Proto-Indo-European to Slavic". *Journal of Indo-European studies* 22: 91–112. [Reprinted in Kortlandt 2011, 157–76.]

————. (2009) *Baltica and Balto-Slavica*. Amsterdam and New York: Rodopi. [*Leiden Studies in Indo-European*, 16.]

————. (2011) *Selected writings on Slavic and general linguistics*. Leiden: Brill. [*Studies in Slavic and General Linguistics*, 39.]

————. (2016) "The inflexion of the Indo-European *o*-stems in Balto-Slavic". *Baltistica* 51(1): 87–96.

Kümmel, Martin Joachim. (2013) "Zur Endung des Genitivs Plural im Indoiranischen". *Indogermanische Forschungen* 118: 193–211.

Lamprecht, Arnošt. (1987) *Praslovanština*. Brno: Univerzita J. E. Purkyně v Brně. [*Spisy Univerzity J. E. Purkyně v Brně, Filozofická Fakulta*, 266.]

Lehr(-Spławiński), Tadeusz. (1917) *Ze studjów nad akcentem słowiańskim*. Cracow: Gebethner; Warsaw: Gebethner i Wolff. [*Prace Komisji Językowej Akademji Umiejętności w Krakowie*, 1.]

Leskien, August. (1876) *Die Declination im Slavisch-Litauischen und Germanischen*. Leipzig: Hirzel.

————. (1919) *Grammatik der altbulgarischen (altkirchenslavischen) Sprache*. 2. und 3. Auflage. Heidelberg: Winter.

————. (1955) *Handbuch der altbulgarischen (altkirchenslavischen) Sprache. Grammatik – Texte – Glossar*. 7. Auflage. Heidelberg: Winter.

Liewehr, Ferdinand. (1955) *Slawische Sprachwissenschaft in Einzeldarstellung*. Vienna: Rohrer.

Ljapunov, B. M. (1905) *Formy sklonenīja v" staroslavjanskom" jazykě*. 1. *Sklonenīe imen"*. Odessa: «Èkonomičeskaja» tipografīja.

Lohmann, J. F. (1930) "Zum slavischen Gen. Sing. der *ā*-Deklination". *Zeitschrift für slavische Philologie* 7(3–4): 372–77.

Lunt, Horace G. (2001) *Old Church Slavonic grammar*. 7th rev. ed. Berlin and New York: Mouton de Gruyter.

Majer, Marek. (2011) "PIE **so*, **seh₂*, **tod*/PSl. **tъ*, **ta*, **to* and the development of PIE word-final **-os* in Proto-Slavic". Thomas Krisch and Thomas Lindner, eds. *Indogermanistik und Linguistik im Dialog. Akten der XIII. Fachtagung der Indogermanischen Gesellschaft vom 21. bis 27. September 2008 in Salzburg*. Wiesbaden: Reichert, 352–60.

Majer, Marek. (2017) *The Caland system in the north: Archaism and innovation in property-concept/state morphology in Balto-Slavic*. Ph.D. dissertation, Harvard University.

Mareš, František Václav. (1963) "Vznik a raný vývoj slovanské deklinace". Bohuslav Havránek, ed. *Československé přednášky pro V. mezinárodní sjezd slavistů v Sofii*. Prague: Československá Akademie Věd, 51–69.

————. (1964) "The Proto-Slavic and early Slavic declension system". *Travaux linguistiques de Prague* 1: 163–72. [Reprinted in Mareš 2001, 33–43.]

————. (1969) *Diachronische Phonologie des Ur- und Frühslavischen*. Munich: Sagner. [*Slavistische Beiträge*, 40.]

————. (1999) *Diachronische Phonologie des Ur- und Frühslavischen*. Herausgegeben und editiert von Angela Bergermayer und Georg Holzer, mit zwei von Hanna Vintr aus dem Tschechischen übersetzten sowie drei weiteren Aufsätzen. Frankfurt am Main: Peter Lang. [*Schriften über Sprachen und Texte*, 4.]

————. (2001) *Diachronische Morphologie des Ur- und Frühslavischen*. Hrsg. und redigiert von Angela Bergermayer. Frankfurt am Main: Peter Lang. [*Schriften über Sprachen und Texte*, 5.]

Matasović, Ranko. (2008) *Poredbenopovijesna gramatika hrvatskoga jezika*. Zagreb: Matica Hrvatska.

Mathiassen, Terje. (1989) "Nochmals der Akk. Pl. der *ā*-Stämme im Ostbaltischen". *Baltistica* 25(2): 123–25.

Meillet, Antoine. (1897) *Recherches sur l'emploi du génitif-accusatif en vieux-slave*. Paris: Bouillon.

————. (1902) *Études sur l'étymologie et le vocabulaire du vieux slave*. Paris: Bouillon. [*Bibliothèque de l'École des Hautes Études* 139.]

————. (1914–15) "De quelques finales slaves". *Rocznik Slawistyczny* 7: 1–8.

————. (1934) *Le slave commun*. 2de éd., revue et augmentée avec le concours de André Vaillant. Paris: Honoré Champion. [Nouveau tirage, 1965.]

Meiser, Gerhard. (1998) *Historische Laut- und Formenlehre der lateinischen Sprache*. Darmstadt: Wissenschaftliche Buchgesellschaft.

Mikkola, Jooseppi Julius. (1897) "Baltische etymologien". *Beiträge zur Kunde der indogermanischen Sprachen* 22(3–4): 239–55.

————. (1950) *Urslavische Grammatik. Einführung in das vergleichende Studium der slavischen Sprachen*. III. Teil: *Formenlehre*. Heidelberg: Winter.

Miklosich, Franz. (1876) *Vergleichende Grammatik der slavischen Sprachen*. Dritter Band: *Wortbildungslehre*. Zweite Ausgabe. Vienna: Braumüller.

————. (1879) *Vergleichende Grammatik der slavischen Sprachen*. Erster Band: *Lautlehre*. Zweite Ausgabe. Vienna: Braumüller.

Milewski, Tadeusz. (1932) "Rozwój fonetyczny wygłosu prasłowiańskiego". *Slavia* 11: 1–32, 225–64.

Moszyński, Leszek. (1984) *Wstęp do filologii słowiańskiej*. Warsaw: Państwowe Wydawnictwo Naukowe.

Moszyński, Leszek. (2006) *Wstęp do filologii słowiańskiej*. Wydanie drugie zmienione. Warsaw: Wydawnictwo Naukowe PWN.

Mueller, Friedrich. (1870) "Documents pour l'étude de la morphologie des langues letto-slaves". *Revue de linguistique et de philologie comparée* 4: 261–68.

Newman, Lawrence. (1971) "On reconstructing a third *jat'* in the northern dialects of Common Slavic". *Slavia* 40(3): 325–41.

Nitsch, Kamizierz. (1928) "W sprawie «trzeciego ě»". V. N. Peretc, ed. *Sbornik statej v čest' akademika Alekseja Ivanoviča Sobolevskogo*. (= *Sbornik Otdelenija Russkogo Jazyka i Slovesnosti Akademii Nauk SSSR* 101(3).) Leningrad: Izdatel'stvo Akademii Nauk SSSR, 359–60.

Noha, Miloš. (1924) "Praslovanská dvojice *-ę* || *-ě*". *Listy filologické* 51(4–5): 244–63.

————. (1927) "Zum Alter der urslavischen Nasalvokale". *Zeitschrift für slavische Philologie* 4: 64–68.

Olander, Thomas. (2009) *Balto-Slavic accentual mobility*. Berlin and New York: Mouton de Gruyter. [*Trends in Linguistics, Studies and Monographs*, 199.]

————. (2012) "Proto-Indo-European **-os* in Slavic". *Russian linguistics* 36: 319–41.

————. (2015) *Proto-Slavic inflectional morphology: A comparative handbook*. Leiden and Boston: Brill. [*Brill's Studies in Indo-European Languages & Linguistics*, 14.]

Orr, Robert. (1985) "The effect of the loss of final *-s on the evolution of Common Slavonic nominal morphology". *New Zealand Slavonic journal* [1985]: 17–32.

————. (1986) "The development of declensional endings in nasals in Common Slavonic". *New Zealand Slavonic journal* [1986]: 83–91.

————. (1988) "A phantom sound-change: CS **-ŏm# >*-ŭ#". *Canadian Slavonic papers* 30(1): 41–61.

————. (2000) *Common Slavic nominal morphology: A new synthesis*. Bloomington, IN: Slavica Publishers.

Otrębski, Jan. (1921) "Słowiańskie formacje na *-ę* || *-ě* i na *-y* || *-a*". *Prace lingwistyczne ofiarowane J. Baudouinowi de Courtenay*. Cracow: W drukarni Uniwersytetu Jagiellońskiego pod zarządem J. Filipowskiego, 10–16.

Pohl, Heinz Dieter. (1983) "Zur Entwicklung der slavischen Substantivdeklination". *Klagenfurter Beiträge zur Sprachwissenschaft* 9: 28–43.

————. (1985) "Paradigmatischer Ausgleich und Funktionswandel in der Frühgeschichte des slavischen Deklinationssystems". Bernfried Schlerath, ed. *Grammatische Kategorien: Funktion und Geschichte. Akten der VII. Fachtagung der Indogermanischen Gesellschaft, Berlin, 20.–25. Februar 1983*. Wiesbaden: Reichert, 370–83.

Porzeziński (Poržezinskij), Wiktor Jan. (1914) *Sravnitel'naja grammatika slavjanskix" jazykov"*. Vyp. 1. Vvedenie. Obščeslavjanskij jazyk. 2nd edition. Moscow: Kušnerev" i K°, 1916.

Porzeziński (Poržezinskij), Wiktor Jan. (1916) *Sravnitel'naja morfologīja drevneindījskago, grečeskago, latinskago i staroslavjanskago jazykov"*. Čast' pervaja: *Sklonenīe imen"*. Moscow: Kušnerev" i K°.

Rasmussen, Jens Elmegård. (1992) "'Woman' in Indo-European, Celtic, and Germanic: One paradigm or two?". *Copenhagen working papers in linguistics* 2: 37–43. [Reprinted in Rasmussen 1999, 505–11.]

————. (1999) *Selected papers on Indo-European linguistics, with a section on comparative Eskimo linguistics*. 2 vols. Copenhagen: Museum Tusculanum.

Ringe, Don. (2006) *A linguistic history of English*. Volume I: *From Proto-Indo-European to Proto-Germanic*. Oxford: Oxford University Press.

————. (2017) *A linguistic history of English*. Volume I: *From Proto-Indo-European to Proto-Germanic*. Second edition. Oxford: Oxford University Press.

Rix, Helmut. (1986) "Die Endung des Akkusativ Plural commune im Oskischen". Annemarie Etter, ed. *O-o-pe-ro-si: Festschrift für Ernst Risch zum 75. Geburtstag*. Berlin and New York: Walter de Gruyter, 583–97.

Rosenkranz, Bernhard. (1955) *Historische Laut- und Formenlehre des Altbulgarischen (Altkirchenslavischen)*. Heidelberg: Winter.

Sandbach, Edmund. (1925) "Die aksl. Endungen *-y, -'ę* im Gen. Sg. bei den fem. *ā-* bzw. *i̯ā-*Stämmen". *Archiv für slavische Philologie* 39: 133–39.

Šaxmatov, A. A. (1915) *Očerk" drevnějšago perīoda istorīi russkago jazyka*. Petrograd: Tipografīja Imperatorskoj Akademīi Nauk.

Schelesniker, Herbert. (1962) "Slav. *toję, tojě* – ai. *tasyām, tasyāḥ*". *Die Sprache* 8(1): 59–62.

————. (1964) *Beiträge zur historischen Kasusentwicklung des Slavischen*. Graz and Cologne: Böhlau. [Wiener Slavistisches Jahrbuch, Ergänzungsband V.]

————. (1967) "Beiträge zur historischen Kasusentwicklung des Slavischen". *Anzeiger für slavische Philologie* 2: 122–30.

————. (1976) "Der slavische Genitiv auf *-y/-ę* und der awestische Lokativ auf *-ąm*". Heinz Dieter Pohl and Nikolai Salnikow, eds. *Opuscula slavica et linguistica: Festschrift für Alexander Issatschenko*. Klagenfurt: Heyn, 383–91. [*Universität für Bildungswissenschaften, Klagenfurt. Schriftenreihe Sprachwissenschaft*, 1.]

Schenker, Alexander. (1993) "Proto-Slavic". Bernard Comrie and Greville G. Corbett, eds. *The Slavonic languages*. London and New York: Routledge, 60–121.

————. (1995) *The dawn of Slavic: An introduction to Slavic philology*. New Haven: Yale University Press.

Scherer, Wilhelm. (1868) *Zur Geschichte der deutschen Spache*. Berlin: Weidmannsche Buchhandlung.

Schleicher, August. (1852) *Die Formenlehre der kirchenslawischen Sprache, erklärend und vergleichend dargestellt*. Bonn: König.

Schleicher, August. (1861–62) *Compendium der vergleichenden Grammatik der indogermanischen Sprachen*. I. *Lautlehre* (1861). II. *Formenlehre* (1862). Weimar: Böhlau.

————. (1871) *Compendium der vergleichenden Grammatik der indogermanischen Sprachen*. Dritte berichtigte und vermerte Auflage. Weimar: Böhlau.

Schmalstieg, William. (1968) "Slavic morpheme alternants in *ě/ę* and *a/y*". *Slavic and East European journal* 12(1): 44–52.

————. (1971) "Die Entwicklung der *ā*-Deklination im Slavischen". *Zeitschrift für slavische Philologie* 36(1): 130–46.

Schmidt, Gernot. (1985) "Der Genetiv der indogermanischen *ā*-Deklination". Hermann M. Ölberg and Gernot Schmidt, eds. *Sprachwissenschaftliche Forschungen. Festschrift für Johann Knobloch zum 65. Geburtstag am 5. Januar 1984 dargebracht von Freunden und Kollegen*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck, 393–402. [*Innsbrucker Beiträge zur Kulturwissenschaft*, 23.]

Schmidt, Johannes. (1883) "Das suffix des participium perfecti activi". *Zeitschrift für vergleichende Sprachforschung* 26(4): 329–400.

Shevelov, George Y. (1964) *A prehistory of Slavic. The historical phonology of Common Slavic*. Heidelberg: Winter.

Sobolevskij, A. I. (1881) "Izslědovanīja v" oblasti russkoj grammatiki". *Russkīj filologičeskīj věstnik"* 6(1–2): 15–53.

Stang, Christian. (1965) "Indo-européen **gwōm*, **d(i)i̯ēm*". Adam Heinz et al., eds. *Symbolae linguisticae in honorem Georgii Kuryłowicz*. Wrocław: Polska Akademia Nauk, 292–96.

————. (1966) *Vergleichende Grammatik der baltischen Sprachen*. Oslo, Bergen, and Tromsö: Universitetsforlaget.

Streitberg, Wilhelm. (1891) "Der Genetiv Pluralis und die baltisch-slavischen Auslautgesetze". *Indogermanische Forschungen* 1: 259–99.

————. (1894) "Vokaldehnung vor tautosyllabischem *-ns* im Baltischen". *Indogermanische Forschungen* 3: 148–56.

Szemerényi, Oswald. (1970) *Einführung in die vergleichende Sprachwissenschaft*. Darmstadt: Wissenschaftliche Buchgesellschaft.

————. (1990) *Einführung in die vergleichende Sprachwissenschaft*. 4th revised edition. Darmstadt: Wissenschaftliche Buchgesellschaft.

————. (1996) *Introduction to Indo-European linguistics*. Oxford: Clarendon Press; Oxford and New York: Oxford University Press.

Tedesco, Paul. (1951) "Review of Mikkola 1950". *Language* 27(2): 165–77.

Townsend, Charles E. and Laura A. Janda. (1996) *Common and comparative Slavic: Phonology and inflection with special emphasis on Russian, Polish, Czech, Serbo-Croatian, and Bulgarian*. Columbus, OH: Slavica Publishers.

Tremblay, Xavier. (1996) "Zum suffixalen Ablaut o/e in der athematischen Deklination des Indogermansichen". *Die Sprache* 38(1): 31–70.

Trummer, Manfred. (1978) "Zu den slavischen Nasalvokalen und den Partnern *-y* (*-a*) : *-ę* (*-ě*) in der Flexion". *Wiener slavistisches Jahrbuch* 24: 254–61.

Vaillant, André. (1950) *Grammaire comparée des langues slaves*. Tome I: *Phonétique*. Lyon: IAC.

————. (1958) *Grammaire comparée des langues slaves*. Tome II: *Morphologie*. Première partie: *Flexion nominale*. Deuxième partie: *Flexion pronominale*. Lyon: IAC.

Vermeer, Willem R. (1991) "The mysterious North Russian nominative singular ending *-e* and the problem of the reflex of Proto-Indo-European **-os* in Slavic". *Die Welt der Slaven* 36(1–2): 271–95.

————. (1996) "Historical dimensions of Novgorod inflexion". A. A. Gippius, T. M. Nikolaeva, and V. N. Toporov, eds. *Rusistika, Slavistika, Indoevropeistika*: *Sbornik k 60-letiju Andreja Anatol'eviča Zaliznjaka*. Moscow: «Indrik», 41–54.

Viredaz, Rémy. (2009) "Two unrecognized vowel phonemes in Proto-Slavic?". Unpublished ms. Online version available at: https://www.academia.edu/6778543/2009_Two_unrecognized_vowel_phonemes_in_Proto-Slavic_Handout_6.09.2009_.

Vondrák, Václav [Wenzel]. (1898) "Einige Bemerkungen anlässlich Meillet's "Recherches sur l'emploi du génitif-accusatif en vieux-slave", Paris 1897". *Archiv für slavische Philologie* 20: 325–42.

————. (1904) "K výkladu některých padů slovanské deklinace". V. I. Lamanskij, ed. *Stat'i po slavjanoveděniju*. Vyp. 1. St. Petersburg: Tipografīja Imperatorskoj Akademīi Nauk", 184–93.

————. (1906) *Vergleichende slavische Grammatik*. I. Band: *Lautlehre und Stammbildungslehre*. Göttingen: Vandenhoeck und Ruprecht.

————. (1908) *Vergleichende slavische Grammatik*. II. Band: *Formenlehre und Syntax*. Göttingen: Vandenhoeck und Ruprecht.

————. (1924) *Vergleichende slavische Grammatik*. Zweite stark vermehrte und verbesserte Auflage. I. Band: *Lautlehre und Stammbildungslehre*. Göttingen: Vandenhoeck und Ruprecht.

————. (1928) *Vergleichende slavische Grammatik*. Zweite stark vermehrte und verbesserte Auflage. II. Band: *Formenlehre und Syntax*. Göttingen: Vandenhoeck und Ruprecht.

Weiss, Michael. (2011) *Outline of the historical and comparative grammar of Latin*. 2nd, corrected printing. Ann Arbor, MI: Beech Stave.

van Wijk, Nicolaas. (1916) "*-ę* und *-ě* im Akk. Plur. der *jo*-Stämme und im Gen. Sing., Nom. Akk. Plur. der *jā*-Stämme". *Archiv für slavische Philologie* 36: 460–64.

Witczak, Krzysztof Tomasz. (1998) "Dwa terminy prasłowiańskie w ujęciu diachronicznym (1. **zъly* > **zъlъva* 'siostra męża, szwagierka', 2. **plo* 'otwarta przestrzeń wodna')". *Slavia Occidentalis* 55: 133–38.

Zaliznjak, A. A. (1988) "Drevnenovgorodskij dialekt i problemy dialektnogo členenija pozdnego praslavjanskogo jazyka". *Slavjanskoe jazykoznanie: X. meždunarodnyj s"ezd slavistov. Sofia, 1988 g. Doklady rossijskoj delegacii*. Moscow: «Nauka», 164–77.

————. (2004) *Drevnenovgorodskij dialekt*. 2-e izdanie, pererabotannoe s učetom materiala naxodok 1995–2003 gg. Moscow: Jazyki slavjanskoj kul'tury.

Zinkevičius, Zigmas. (1980) *Lietuvių kalbos istorinė gramatika*. I: *Įvadas. Istorinė fonetika. Daiktavardžių linksniavimas*. Vilnius: "Mokslas".

Zubatý, Josef. (1893) "Zur Declination der sog. *-iā-* und *-io-*Stämme im Slavischen". *Archiv für slavische Philologie* 15: 493–518.

————. (1896) "Baltische Miszellen". *Indogermanische Forschungen* 6: 269–307.

————. (1897 [1898]) "Über gewisse Genitivendungen des Lettischen, Slavischen und Altindischen". *Věstník Královské české společnosti nauk* 1897, č. XVII, 1–27.

Zucha, Ivo. (1986) "Zu den Endungen *-y, -ę, -ě* im G. Sg. und N. Pl. der slawischen Deklination und zu den Nasalvokalen". *Wiener slavistisches Jahrbuch* 32: 133–37.

Department of Older Germanic Languages
Faculty of English
Adam Mickiewicz University in Poznań
al. Niepodległości 4
61-874 Poznań, Poland
ronald.kim@yahoo.com

Clitic Pseudo-Climbing out of Dependent Infinitive Phrases in Serbian*

Jasmina Milićević

Abstract: The paper describes the linear placement of Serbian clitics syntactically depending on an infinitive in the role of verbal, nominal, or adjectival complement (for short, infinitive clitics). Two linear placement options are in principle available for the infinitive clitics: local placement (in the second linear position [2P] of the infinitive phrase, where they form a cluster of their own) and clause-level placement (in the 2P of the clause, where they join clause-level clitics). Syntactic and prosodic conditions licensing these placement options are discussed and the corresponding clitic linearization rules proposed within a Meaning-Text dependency-based approach to syntax (Mel'čuk 1988, 2013; Polguère and Mel'čuk 2009). It is shown that clause-level placement of 2P clitics (unlike that of ad-verbal clitics of the Romance type) does not require the change of the syntactic governor of the clitics and should therefore be set apart from genuine clitic climbing. This is why the corresponding phenomenon is termed "clitic pseudo-climbing".

1. Overview of the Problem

The paper deals with a specific word-order phenomenon in Serbian: linear placement of the clitics that syntactically depend on an infinitive subordinated to another lexeme—typically a verb, but also a noun or an adjective. For ease of reference, these clitics—the reflexive marker SE and clitic forms of personal pronouns (in the genitive, dative, and accusative)—will be called infinitive clitics.[1]

* This paper is an extended version of a talk given at the 2016 Toronto meeting of the Slavic Linguistic Society.

[1] I will assume familiarity with such basic notions as *clitic*, *clitic cluster*, *host of a clitic*, etc. Nevertheless, explanations deemed necessary will be provided. For an overview of Serbian clitics, which are second position clitics and include, in addition to the above-mentioned items, clitic forms of auxiliary verbs, an interrogative, and an emphatic particle, see, for instance, Spencer 1991: 351–58 and Franks and Holloway King 2000: 17–31.

Among the dative clitic pronouns, only genuine object clitics will be considered, i.e., those realizing inherent syntactic actants (a.k.a., in other frameworks, syntactic arguments) foreseen in the government pattern of their governors. Possessive and ethical dative clitics are excluded from the discussion.

While clause-level clitics are gathered into a single cluster and placed, roughly speaking, in the second linear position [2P] in the clause,[2] infinitive clitics have, in principle, two basic placement options:

1) Non-local placement: clause-level placement, in a single cluster with clause-level clitics (if present).
2) Local placement: placement in the 2P within the infinitive phrase, with clause-level clitics either joining the cluster of infinitive clitics or forming a separate cluster outside of the infinitive phrase.

In the linguistic literature, non-local placement of the infinitive clitics is commonly viewed as the result of clitic climbing out of the infinitive phrase into the "matrix clause". However, in the Meaning-Text dependency framework (Mel'čuk 1988, 2013; Polguère and Mel'čuk 2009), the term *climbing* turns out to be a misnomer when applied to 2P clitics. As we shall see below (Section 3), 2P clitics do not "climb" in the sense that ad-verbal clitics in Romance languages do, i.e., they do not change the syntactic governor. At most, 2P clitics display what could be called "climbing effects". For this reason, this particular clitic placement option will be referred to as clitic pseudo-climbing [CPC].

Whether infinitive clitics are placed non-locally or locally depends on a combination of syntactic and prosodic factors (to be discussed in more detail in Section 2)—in the first place on the linear position/syntactic role of the infinitive phrase and the "prosodic weight" of the infinitive phrase/the rest of the clause. Together, these factors determine what could be called the *degree of integration* of the infinitive phrase into the clause. What does it mean for an infinitive phrase to be well integrated? On the one hand, the infinitive phrase is semantically and syntactically closely linked with its governor, which is the case, for instance, if the governor is a verb of particular type (the future auxiliary, a modal, etc.). On the other hand, the infinitive phrase is not prosodically independent from the rest of the clause, i.e., it does not bear its own prosodic contour and is not set off by a prosodic break; this happens, for example, if the infinitive phrase contains only an infinitive verb and clitics, which makes it prosodically light. Roughly speaking, a high degree of integration of the infinitive phrase into the clause results in the construction of a single clitic clus-

[2] "Roughly speaking" because constituents of certain types do not qualify as hosts for the clitic cluster and some that do can be skipped, so that the cluster sometimes "lands" in a position further removed from the left edge of the clause (see Halpern 1995: 67ff and Milićević 2009a: 248ff).

ter comprising both the clause-level and the infinitive clitics. This happens of course in the situation of non-local placement but also in some instances of local placement.

The sentences in (1a) through (1d) below provide some preliminary examples (clitics are in boldface). The surface-syntactic governor of the infinitive verb (V_{INF}) is boxed. The double vertical bar "||" symbolizes an optional prosodic break. The infinitive phrases are in square brackets "[…]". Clitics that may climb but do not have to are in angle brackets "<…>".[3]

While processing the examples, the reader should be aware of the following two facts:

1) Serbian auxiliary verbs are clausal heads (in the syntactic structure of the clause) independently of their tonicity status: clitic or full (i.e., stressed); see Milićević (1999). Negative forms of auxiliary verbs (*nisam* 'I am not', *ne bih* 'I would not', *neću* 'I will not', etc.) are always stressed.
2) Serbian is a pro-drop language: in communicatively neutral environments, the pronominal syntactic subject of the clause is elided on the surface (i.e., it does not have a phonetic realization but is present in the syntactic structure of the clause).

(1) a. Ne boxed{moraš} **mi** [ništa *<**mi**> reći], svoju prošlost,
not must to.me nothing to.tell your past,
svoje ime … (Song Lyrics)
your name
'You do not need to tell me anything, your past, your name …'

b. Baš **bih** <**te**> boxed{volela} || [videti <**te**> s kratkom
really COND you$_{SG}$ like$_{PTCP}$ to see with short
kosom]. (WWW)
hair
'I would really like to see you with short hair.'

[3] Sources of examples: *Korpus savremenog srpskog jezika* (Corpus of Contemporary Serbian), hereafter SerbCor (www.korpus.matf.bg.ac.rs), WWW, linguistic literature; two examples are taken from Nušić (1924); the remaining are my examples. When examples from a cited source include an alternative version, such as all four examples in (1), the judgment of the alternative version is my own. Most of the WWW and my examples have been checked with other native speakers who overwhelmingly corroborated my acceptability judgments.

(1) c. [Pohapsiti **bi** **ih**] trebalo *<**bi**>.
to.arrest COND them having.been.necessary
(Adapted from Popović 1997: 350)
'Arrested is what they should be.'

d. [Tražiti **ga** *<**je**> u mraku] ?**je** teško <|| teško **je**>.
to.seek him is in darkness difficult
(Adapted from Browne 1975: 275)
'Looking for him in the dark is difficult.'

In examples (1a–b), the infinitive phrase is not clause-initial, which represents the unmarked word order. V_{INF} functions as the object of a restructuring verb (see Section 2.1.2): a modal verb and an emotional-state verb, respectively. Non-local placement of the infinitive clitics, i.e., CPC, is obligatory in (1a), and optional in (1b). In the latter case, the accusative clitic *te* 'you$_{SG}$' may be placed locally if a prosodic break is made before the infinitive phrase, which is possible because of the prosodic heaviness (roughly, the number of syllables) of both the infinitive phrase and the rest of the clause.

In examples (1c–d), the infinitive phrase is clause-initial: this is the marked word order. In (1c), V_{INF} is the object of a restructuring verb. In (1d) it is the subject of the copula. In both cases, only local placement is available for the infinitive clitics, i.e., there is no CPC; clause level clitics must be placed together with the infinitive clitics in (1c), and separately in (1d). Moreover, in (1d), at least for some speakers, there is a preference for the clause-level clitic cluster to not follow the infinitive phrase immediately;[4] rather, the infinitive phrase is prosodically set off from the rest of the clause and skipped (cf. footnote 3), i.e., not counted as a possible host for the copula.

There are no semantic differences between sentences where the infinitive clitics are placed locally vs. non-locally, since clitics are not content words. There are no communicative differences either, because linear placement of the clitics, unlike that of full-fledged clause elements, is not flexible enough to be used as a means of expressing communicative distinctions. As for rhetorical (stylistic) differences, non-local vs. local placement of the infinitive clitics reflects style/register, as well as dialectal/idiolectal preferences.

Overall, CPC is not a prominent phenomenon in Serbian. On the one hand, it represents a "micro-variation" in word order, affecting a part of an already "closed system" that clitics represent. On the other hand, the frequency of CPC in texts must be rather low due to a restricted use of the infinitive in Serbian, which in most of its syntactic roles can be and preferably is replaced

[4] Since a clitic cluster can be construed as a set of clitics and a set can consist of only one element, I will be freely speaking about clusters containing a single clitic, like the one in sentence (1d).

by the complement clause—a (finite) *da*-clause (*da* 'that' is a complementizer/conjunction).[5] Nevertheless, the phenomenon deserves attention from both a descriptive and a theoretical viewpoint: there are instances where the infinitive is used quite naturally in Serbian, and where, moreover, CPC of its clitics is obligatory; and CPC is interesting because its nature has not been fully explained.

Another sub-finite domain in which clitics can be placed is a phrase headed by a converb (a non-finite verb expressing relative tense: simultaneity or anteriority with respect to the MV). Converb phrase clitics do not exhibit CPC.

Finally, let us note that CPC out of complement *da*-clauses was marginally possible in older stages of the language but is virtually nonexistent in standard modern Serbian and will not be considered in this paper.[6]

---

[5] On the use of infinitive in the Balkan Sprachbund, in particular in Serbian, see Mišeska-Tomić 2006.

[6] In the linguistic literature there has been some controversy over the possibility of CPC out of a complement *da*-clause; see Jurkiewicz-Rohrbacher, Kolaković, and Hansen 2017a: 50 for an overview. My own intuition is that this variety of CPC is only marginally possible in some registers (journalistic, literary), and perhaps more common in Bosnian than Serbian. The marginality of the phenomenon seems to be corroborated in Jurkiewicz-Rohrbacher, Kolaković, and Hansen 2017b, where extremely low frequencies of CPC out of *da*-clauses were reported for a significantly large corpus of texts.

Here are some examples of CPC out of complement clauses found in the linguistic literature. Grammaticality judgments are mine (in the original texts, sentences in (i-a) were judged in the same way, while the remaining two were evaluated as grammatical). I find (i-c), with a "partial CPC", severely ungrammatical.

(i) a. Milan ?**ga**$_i$ želi [da $\_\_{}_i$ vidi]. (Progovac 1993: 134)
M. him wants that$_{(CONJ)}$ sees
'M. wants to see him.'

vs.

Milan ***ga**$_i$ kaže [da $\_\_{}_i$ vidi].
says
'M. says that he sees him.'

b. To ?**mu**$_i$ ne želim [da $\_\_{}_i$ posudim]. (Aljović 2006: 1)
this to.him not want that$_{(CONJ)}$ lend
'I do not want to lend this to him'.

c. Marija ***mu**$_i$ želi [da $\_\_{}_i$ **ga** predstavi].
M. to.him wants that$_{(CONJ)}$ him introduces
'M. wants to introduce him to him.' (Stjepanović 1998, in Franks 2010: 21)

The contexts in which CPC out of *da*-clause is marginally possible (called *subjunctive-like* in Progovac 1993) are the same ones where it can happen out of infinitive

While publications on Serbian clitics (more generally, clitics in Bosnian, Croatian, and Serbian [BCS]) are plenty, CPC as such has received relatively little attention. A detailed, non-formal description of CPC in literary Serbo-Croatian is found in Popović 1997: 342–50. Browne's (1975 / 2004) seminal paper contains the first description of the phenomenon in standard Serbo-Croatian within the generative framework (273–76). Most of the subsequent work on CPC has also been done in phrase-structure based approaches to syntax. Discussion and examples of CPC in BCS are found, for instance, in Ćavar and Wilder 1999; Franks 2010; Bošković 2004, 2015; Progovac 2005: 146ff; Zimmerling 2011; Zimmerling and Kosta 2013. Papers focused on or fully dedicated to the topic include Stjepanović 2004, Caink 2004, Aljović 2006 and Jurkiewicz-Rohrbacher, Kolaković, and Hansen 2017a/b. As for dependency-oriented studies of BCS clitics *tout court*, I am aware only of Čamdžić and Hudson 2002 and Milićević 2009a, neither of which speaks about CPC at any length.[7]

---

phrases, namely the governor of the *da*-clause has to be a verb of particular type (a restructuring verb, as will be explained later), here *želeti* 'to want'. With a non-restructuring matrix verb, such as *kazati* 'to say', CPC is clearly unacceptable. Such verbs cannot take an infinitive phrase as a complement.

Examples of CPC out of *da*-clauses in literary Serbian, which sound dated to the contemporary ear, are given in (ii-a–b). Sentence (ii-c) is an example from the press and (ii-d) was spontaneously produced in a conversation I had with a speaker of Bosnian.

(ii) a. Kad **bi** **ga**$_i$ žena počela […] [da $\_\_{}_i$ teši] …
when COND him wife start$_{PTCP}$ that$_{(CONJ)}$ consoles
'When his wife would start to console him …' (In Simić 1999: 341)

b. jer **ga**$_i$ njegova bivša supruga nije umela dovoljno
for him his ex wife not.is know$_{PTCP}$ enough
[da $\_\_{}_i$ razume] (Nušić 1924: 2)
that$_{(CONJ)}$ understands
'because his ex-wife was not quite able to understand him'

c. To **im**$_i$ Vučić ne sme [da $\_\_{}_i$ dozvoli].
that to.them V. not may that$_{(CONJ)}$ allows
(In Jurkiewicz-Rohrbacher, Kolaković, and Hansen 2017a: 59)
'That is something V. should not allow them (to do).'

d. Sve **ih**$_i$ **je** htio [da $\_\_{}_i$ vidi]. (PC, non elicited)
all them is want$_{PTCP}$ that$_{(CONJ)}$ sees
'He wanted to see all of them.'

Note the fronting of the direct object of the embedded finite verb in (i-b) and (ii-c) and of the quantifier in (ii-d). These may be factors that make CPC out of a *da*-clause somewhat more acceptable. Whether object/circumstantial extraction have the same effect on CPC out of the infinitive phrase remains to be determined.

[7] For clitic (pseudo-)climbing in Slovenian, the only South Slavic language besides BCS that has it, see Golden 2003 and Marušič 2008. An overview of clitic (pseudo-)

I will propose sample linearization rules for infinitive clitics, in particular those covering CPC out of infinitive phrases. This will be done within a Meaning-Text dependency syntax approach (Mel'čuk 1988, 2013; Polguère and Mel'čuk 2009) and from the viewpoint of linguistic synthesis (= speech production, as opposed to analysis or speech comprehension).

The rest of the paper is structured as follows: Section 2 provides an informal description of the linearization of the infinitive clitics, Section 3 presents linearization rules for the infinitive clitics, and Section 4 concludes the paper.

Before we proceed, a word of caution is in order. Grammaticality judgment of sentences containing infinitive clitics is difficult. First, due to differences in the infinitive clitic placement across BCS and across different registers (journalistic, literary, spoken, etc.) to which most speakers are exposed (through media, etc.), contamination can happen quite easily. Second, the intuition of speakers of Serbian with regard to infinitive clitic placement may be further impaired by the fact, already mentioned, that in modern Serbian (in contrast especially to Croatian) the range of contexts in which the infinitive is used is more limited. These factors explain, at least to some extent, the disagreements, not infrequently found in the literature, about the acceptability of sentences containing, in particular, non-locally positioned infinitive clitics (cf. comments made to this effect in Jurkiewicz-Rohrbacher, Kolaković, and Hansen, 2017a: 50ff).

2. An Informal Description of the Linearization of Infinitive Clitics

We start by discussing the general lexical and syntactic conditions under which CPC out of infinitive phrases is possible—classes of lexemes that can govern a V_{INF} in the syntactic structure of the clause and syntactic roles that a V_{INF} can have (2.1). Then follows a description of specific factors governing non-local vs. local placement of infinitive clitics (2.2).

2.1. Governors and Syntactic Roles of V_{INF}

Classes of lexemes that can govern a V_{INF} in Serbian and the corresponding syntactic roles of V_{INF} are summarized in Table 1 on the following page.

climbing in the Balkan languages can be found in Krapova and Turano 2015.

Table 1. Governors of V_{INF} and syntactic roles of V_{INF}

	Governor of V_{INF}	**Syntactic role of V_{INF}**
1.	*hteti*$_{(AUX)FIN}$ 'will'	Lexical Part of main verb
2.	Semantically full V	Object
3.	*biti*$_{(COPULA/AUX)FIN}$ 'to be'	Subject
4.	N/ADJ	Object
5.	*biti*$_{(COPULA)}$	Copular attribute

The verb governing the V_{INF} (numbers 1, 2, 3, and 5 in the table above), i.e., the (absolute) syntactic head of the clause, will be referred to as the main verb (MV). Thus, a MV can be an auxiliary verb, a copular verb, or a semantically full verb. (This departs from the usage in which *main verb* is synonymous with *lexical* or *semantically full verb*.)

2.1.1. V_{INF} as the Lexical Part of MV in the Future Tense

In the future-tense forms, the infinitive is the lexical part of the main verb (not an object), and the main verb itself is a clitic. In this case, the infinitive clitics always form a single cluster with the clitic future marker and must be placed in the exact same way as the latter; this is shown in (2a).

(2) a. (i) Nekako **ću** **im** **se** oduži+ti.
somehow FUT.1SG to.them REFL to.repay
'I'll repay them (for their kindness) somehow.'
(ii) *Nekako **ću** oduži+ti **im se**.
(iii) Oduži=**ću** **im se** nekako.

Compare (2a) with (2b), where *odužiti (se)* is the object of *morati*, and where the clitics of the former can be placed both locally and non-locally (in this particular context, because of the prosodic heaviness of the infinitive phrase).

(2) b. Mora=**ću** <**im** **se**> [nekako <**im se**> odužiti].
must FUT.1SG to.them REFL somehow to.repay
'I will have to repay them somehow (for their kindness).'[8]

[8] In (2a-iii) and (2b) the infinitive suffix *-ti* of the verbs *oduži+ti (se)* and *mora+ti* is truncated before the future tense marker, which in this context behaves more like an affix than like a word form. In spite of appearances, *ću* is the main verb (i.e., the syntactic

Due to a close semantic and syntactic link between V_{INF} and the future auxiliary, the infinitive is strongly integrated into the clause and not perceived as an independent clause element. This is corroborated by the fact that clauses where the infinitive as the lexical part of the MV occupies the initial position, as in (2a-iii), are less marked than those in which the infinitive phrase functioning as an object or the subject is clause-initial.

2.1.2. V_{INF} as the Object of a Semantically Full Verb

Earlier we used the term restructuring verbs (Rizzi 1982; Monachesi 1998; Dobnik and Cooper 2016), familiar from the literature on clitic climbing in Romance languages,[9] to characterize Serbian verbs that, as an alternative to a *da*-clause, take a V_{INF} as complement (direct, indirect, or oblique object). According to another well-known syntactic classification, these verbs fall into raising, subject-control, and object-control verbs (Nišida 2012; Jurkiewicz-Rohrbacher, Kolaković, and Hansen, 2017b). Semantically, restructuring verbs belong to several classes, as indicated in Table 2 on the following page (cf. Piper et al. 2005: 470ff; Mišeska-Tomić 2006: 414–16, 484–85; Mrazovac and Vukadinović 2009: 130ff).

I call *quasi-modals* the verbs which, like genuine modals, have a modal meaning and are semantically incomplete (i.e., require an infinitive or a clausal complement) but which, unlike the latter, are impersonal (i.e., have a 3p neuter syntactic subject that has no phonetic realization).

Causative and motion verbs can also take V_{INF} as an object; cf., respectively, *Dao* <ih> *je* [*zatvori*i <ih>] 'He had them put in jail' and *Došao sam* <te> [*nešto* <**te**> *pitati*] 'I came to ask you something'. Such constructions are dated in Serbian and more representative of Croatian and Bosnian. As can be seen from the examples, both verb types allow for the CPC of the clitics depending on their infinitive complement.

V_{INF} is most naturally used as a complement of (quasi-)modal, phasic, and emotional state verbs. Its use with verbs from other classes, especially if they

head of its clause) and the future tense forms are analytic, just as those of the compound past or the conditional. (On the morphological status of future-tense markers in Serbian, see Milićević 2009b and references herein.)

[9] The term was introduced to refer to verbs that behave as a class with respect to clitic climbing and some other phenomena, such as "long object preposing" and "unbounded tough movement". To account for their properties, it was suggested that these verbs undergo an optional restructuring rule that makes it possible to analyze the matrix verb and its embedded infinitive as a verbal complex, or as it is sometimes called, a complex predicate. I use the term only as a convenient label, since in the present framework it is not necessary to invoke any such rule in order to account for the corresponding facts.

Table 2. Serbian verbs which take an infinitive complement

Modal:	*moći* 'can', *morati* 'must', *hteti* 'will', *smeti* 'dare'	Raising Verbs
Quasi-modal:	*trebati*$_{V.IMPERS}$ 'be necessary', *valjati*$_{V.IMPERS}$ 'be needed'	
Phasic:	*početi/stati* 'start', *nastaviti* 'continue', *prestati* 'stop'	
Emotional state:	*voleti* 'like', *želeti* 'want', *bojati se* 'fear',	Subject Control Verbs
Mental state/act:	*mrzeti*$_{V.IMPERS}$ 'be hateful to'	
	nameravati 'intend', *umeti* 'know how', *običavati* 'have habit', *zaboraviti* 'forget', *odlučiti* 'decide'	
(Speech) act:	*obećati* 'promise', *usuditi se* 'dare', *odvažiti se* 'venture', *nastojati* 'strive', *pokušati* 'try', *uspeti* 'succeed'	
	narediti 'direct', *zabraniti* 'forbid', *zamoliti* 'ask'; *dozvoliti* 'allow', *pustiti* 'let', *naterati* 'force', *primorati* 'compel'; *pomoći* 'help'	Object Control Verbs

are object control verbs, has a "Croatian feel" and *da*-complementation is generally preferred.

In the remainder of this section, only the unmarked word-order situations are considered, i.e., those in which the infinitive phrase is not clause-initial and both non-local and local placement options are in principle available for the infinitive clitics.

Without prosodic factors intervening, CPC out of the infinitive complements of most restructuring verbs is obligatory. This is especially true for modal and phasic verbs—understandably enough, since they are more closely related to their governor and similar in this respect to auxiliary verbs. However, even with these verbs, prosodic factors may allow for the local placement of infinitive clitics.

Examples in (3) illustrate CPC out of infinitive phrases governed by (quasi-)modal, phasic, and emotional-state verbs: CPC is obligatory in (3a–b) and optional in the remaining examples, which feature prosodically heavy infinitive phrases. Notice in particular the contrast between (3b) and (3d), where with the same verb governing V_{INF}, different linear placing requirements hold for the infinitive clitics.

(3) a. Ekstra kvalitetom <**se**> mogu <**se**> [pohvaliti *<**se**>] i
with extra quality REFL can boast and
pekare … (SerbCor)
bakeries
'Extraordinary quality can be boasted as well by the bakeries …'

b. Nisam <**vam**> to smeo [reći *<**vam**>]! (SerbCor)
not.am to.you$_{PL}$ that$_{(PRON)}$ dare$_{PTCP}$ to.tell
'I shouldn't have told you that!'

c. … Buskeros <**se**> ponovo poče [smejati <**se**> kao lud].
… B. REFL again started to.laugh like crazy
'… B. started laughing his head off again.'

d. … a poreznici <**mu**> ne bi smeli [suviše
and tax collectors to.him not COND dare$_{PTCP}$ too.much
<**mu**> dosađivati]. (SerbCor)
to bother
'… and tax collectors shouldn't be bothering him too much.'

e. … treba <**je**> samo [umeti <**je**> [pročitati ?<**je**>]]
is.necessary it$_{FEM}$ only to.know to.read (SerbCor)
'it is only necessary to know how to read it' (*je* refers to book, a feminine noun.)

The verb governing V_{INF} can be finite, as in (3a) and (3c), or non-finite: a participle ((3b) and (3d)) or another infinitive (3e). While there is no theoretical limit to the number of stacked infinitives, in practice, having more than two is unusual.[10] Note the multiple landing sites for the clitics in (3a) and (3e).

Instances where CPC out of infinitive phrases embedded under restructuring verbs is problematic or blocked seem to involve specific reflexive verbs; cf.:

[10] Cf. the following constructed example from Slovenian (Marušič 2008), with three embedded infinitives:

(i) On <jo> je hotel <jo> [nehati <jo> [hoteti [videvati <jo>
he her is want$_{PTCP}$ to.not.want to.want to.see
vsak dan]]].
every day
'He wanted to not want to want to see her every day.'

Such a sentence is possible in Serbian, as well, but it would not be produced spontaneously. For some corpus examples of stacked infinitives in BCS, see Hansen, Kolaković, and Jurkiewicz-Rohrbacher 2018.

(4) a. (i) Nisam ?**mu** **se** ~~**se**~~ bojao
not.am to.him REFL REFL fear$_{(V.REFL)PTCP}$
[suprotstaviti]. (Adapted from Popović 1997: 345)
to.oppose$_{V.REFL}$
'I wasn't afraid to oppose him.'

(ii) Nisam **se** bojao [suprotstaviti **mu se**].

b. (i) … jer **me** ***se** ~~**je**~~ mrzelo [oblačiti]
for me REFL is having.been.hateful to.dress$_{(V.REFL)}$
(Adapted from Popović 1997: 345)
'… for I didn't feel like dressing'

(ii) … jer **me je** mrzelo [oblačiti **se**]

c. Pomozi **mi** ?<**se**> [setiti <**se**>].
help to.me REFL to.recall$_{(V.REFL)}$
'Help me to recall (something to memory).'

d. Starica **mi** *<**se**> pomogne || [uzdići <**se**>
old.woman to.me REFL helped to.stand.up$_{(V.REFL)}$
na noge]. (Popović 1997: 345)
on feet
'The old woman helped me get back on my feet.'[11]

The ban on CPC holds in the context of two subject-control verbs (4a–b), an object-control verb (4c–d), and regardless of prosodic factors ((4c) vs. (4d)). This makes me think that the culprit is the reflexive nature of the verbs involved. If this is correct, then all such verbs need to be explicitly marked as not allowing for the CPC of their clitics in this specific syntactic environment. For cases where CPC needs to be precluded because it results in undesirable or unacceptable clitic combinations, see 2.2.3 below.

2.1.3. V_{INF} as the Subject of a Finite Copula or Auxiliary Verb

In (5a), the infinitive is the syntactic subject of the copula, while in (5b) it functions as the subject of the auxiliary *biti* 'to be' forming the compound past with the copula. (The role of the copular attribute is filled by the adjective *zanimljivo* 'interesting' in (5a) and the noun *žalost* 'sorrow' in (5b).)

[11] In (4a), both the V_{INF} and the governing verbs are reflexive; in other examples in (4), only the V_{INF} is. Note the haplology of one SE in (4a-i) and the deletion of je from the sequence *[se je] in (4b-i), virtually obligatory in Serbian (see Footnotes 16 and 17).

(5) a. (Emisije radi profesionalno) i zanimljivo **<ih>**
(He creates the programs like a pro) and interesting them
je [pratiti **<ih>**]. (WWW)
is to.follow
'and it is interesting to follow them.'

b. (ali je taj globus izgledao tako bedno) da **<ga>** **je** žalost
(but that globe looked so shabby) that it is sorrow
bila [pogledati **<ga>**]. (Nušić 1924: 22)
be$_{(V.COP)PTCP}$ to.see
'that it made one said just to look at it.'

Since a V_{INF} functioning as the subject is easy to set off prosodically from the rest of the clause, its clitics can be placed either non-locally or locally. (But if this per se legitimate non-local placement results in unacceptable clitic combinations, it must be prevented; see 2.2.3 below.)

2.1.4. V_{INF} as the Oblique Object of N/ADJ

A small number of nouns, such as *želja* 'wish', *volja* will', *namera* 'intention', *smisao* 'purpose', and *razlog* 'reason', take V_{INF} as an oblique object; see (6a).[12]

Adjectives taking V_{INF} as an oblique object mostly denote physical and mental dispositions, for example, *kadar* 'able', *sposoban* 'capable', *voljan* 'willing', *raspoložen* 'disposed', *gotov* 'decided', *spreman* 'ready', *rad* 'eager', *oran* 'enthusiastic', *dužan* 'bound', *nameran* 'intent (on)', *saglasan* 'agreeing'; see (6b).

(6) a. (i) Nemam **<se>** nameru [ni sa kim
not.have REFL intention not with anyone
<se> prepucavati]. (WWW)
to.argue
'I have no intention to argue with anyone.'

(ii) … a ima **<je>** želju [pročitati **<je>**].
and has it$_{FEM}$ desire to.read'
'… and he would like to read it.'

[12] Interestingly, this happens only when these nouns are embedded under *imati* 'to have'/*nemati* 'to not have'; cf.: *Ima*/*Nema nameru otići <da ode>* '[S/he] has intention to.leave <that [s/he] leaves>' vs. *Njegova namera *otići <da ode> nije mi poznata* 'His intention to.leave <that [he] leaves> not.is to.me known'. (More generally, the infinitive complementation is possible only when the governing noun is found in a collocation with a light/realization verb. I am grateful to Wayles Browne (p.c.) for pointing this out.)

(6) b. (i) Željni **smo** <**ih**> i mi [videti <**ih**>]. (WWW)
desirous are them and we to.see
'We too would like to see them.'

(ii) Nisam <**ga**> sposoban [rešiti <**ga**>].
not.am it able to.solve
'I am not able to solve it.'

Constructions of this type are not widely used in Serbian (in these contexts, a *da*-clause is much preferred over a V_{INF}). Both non-local and local placement of the infinitive clitics is possible.

2.1.5. V_{INF} as the Copular Attribute

(7) [Voleti svoje susede] jeste (isto što i) [pomagati **im**
to.love one's neighbors $is_{(V.COP)FULL}$ (the same as) to.help them
kad zatreba].
when need.arises
'To love one's neighbors is to help them in need.'

In (7), the infinitive phrase headed by *pomagati* 'help$_{INF}$' appears as the attribute of the copula *biti* 'be$_{INF}$'; this is a minor V_{INF} syntactic role, as the corresponding construction is rarely used. The clitics of an infinitive functioning as the copular attribute are always placed locally.[13]

2.2. Factors Relevant for the Linear Placement of the Infinitive Clitics

As we have already seen, both syntactic and prosodic factors influence the linear placement of the infinitive clitics. In addition, the number/type of clause-level and infinitive clitics involved may act as a constraining factor for non-local placement.

[13] In a synonymous (and more frequently used) construction involving the verb *značiti* 'to mean', where the infinitive appears in the role of object, the infinitive clitics must also be placed locally (understandably enough, since this verb does not belong to the class of restructuring verbs, indicated in Table 2 above, with which non-local placement is in principle allowed). This fact was mentioned (in a different context) in Browne 1987: 171.

2.2.1. Linear Position of the Infinitive Phrase and Syntactic Role of V_{INF}

The possibility of non-local placement of infinitive clitics is determined by the linear position of the infinitive phrase within the clause: clitics can undergo CPC only out of non-clause-initial infinitive phrases. (This is why the non-local placement of infinitive clitics is called *anticipatory* in Popović 1997: 338.) When the infinitive phrase is clause-initial, CPC is blocked altogether.

Possibilities of CPC out of non-clause-initial infinitive phases are determined in large part as a function of the syntactic role of the V_{INF}. Thus, clitics of a V_{INF} that is the lexical part of the main verb in the future tense can only be placed non-locally. If V_{INF} is the attribute of the copula, its clitics are always placed locally. And if V_{INF} appears as the subject of the clause, both non-local and local placement is available for its clitics.

2.2.2. Prosodic Features of Clause Elements

Prosodic factors play an important role in allowing for optional CPC or making it preferable. They are especially important for configurations in which V_{INF} is the object of a semantically full verb. If the infinitive phrase constitutes an intonational phrase (because it is heavy) and a prosodic break can be made before it, local placement of the infinitive clitics becomes possible. This is why with the same verb there can be different CPC options as a function of prosody. Compare (8a–b) with (1b), repeated here as (8c):

(8) a. Baš **bih** **te** volela [videti].
really COND you$_{SG}$ like$_{PTCP}$ to.see
'I would really like to see you.'

b. Baš **bih** volela || [videti ?**te**].

c. Baš **bih** volela || [videti **te** s kratkom kosom].
with short hair

The same phenomenon is illustrated by the following example (sentences (9a–c) repeat (3e) from above):

(9) a. Treba **je** samo umeti pročitati.
is.necessary it$_{FEM}$ only to.know to.read
'It is only necessary to know how to read it.'

b. Treba samo || [umeti **je** [pročitati]].
is.necessary only to.know it$_{FEM}$ to.read

(9) c. Treba samo [umeti || [pročitati ?**je**]].
is.necessary only to.know to.read it$_{FEM}$

d. Treba samo [umeti || [pročitati **je** na pravi način]].
in proper manner

The dubious acceptability of sentences (8b) and (9c) may be also due to the Final Position Effect (Milićević 2009a: 250–51). While generally speaking clitics tend to stay away from the right edge of the clause, this is not an absolute requirement, and the clitics in (11) below are perfectly fine sitting in the clause-final position.

Thus CPC correlates with a high level of semantic, syntactic, and prosodic integration of the infinitive phrase into the clause; cf. Popović 1997: 345ff "[in the situations of non-local placement] the infinitive is informationally closely linked to the [main] verb and carries the same sentential stress", and "[in local placement] the infinitive phrase [appears] as an informational and articulatory unit".

Table 3 below summarizes the linear placement possibilities existing for the infinitive clitics.

Table 3. CPC options available to infinitive clitics

Syntactic Role of V_{INF}	**CPC of the clitics of V_{INF}**	
	Inf. Phrase NOT Clause-Initial [the unmarked case]	Inf. Phrase Clause-Initial
1. Lexical Part of MV	100% OBLIGATORY	BLOCKED
2. Object of V	OBLIGATORY or OPTIONAL	
3. Subject of V	OPTIONAL	
4. Object of N/ADJ	OPTIONAL	
5. Copular Attribute	100% BLOCKED	N/A[14]

2.2.3. Type and Number of Clitics Involved

Optional CPC out of any of the configurations indicated in Table 3 can be precluded for specific clitic combinations if these violate constraints on the constitution of the clitic cluster. Some of the constraints arise only for mixed

[14] N/A because a copular attribute is never clause-initial.

clusters (e.g., ambiguity/high number of clitics; case incompatibility), and some are independent of CPC (e.g., person-case constraint). They are modeled as filter rules, some of which are presented in this section (2.2.3.1 through 2.2.3.4) and in 3.2.3 below.

As mentioned earlier, some constraints on CPC arise for specific verbs, which need to be marked accordingly in the lexicon; cf. example (4) above and the related remarks.

2.2.3.1. CPC Resulting in Ambiguity

(10) Ne vredi **mu** **ga** [prodavati].
not is.worth to.him it to.sell

The indirect object clitic *mu* in (10) can be interpreted as depending on the MV [*vredi*-indir.objectival→*mu*] or as depending on the V_{INF} ([*prodavati*-indir. objectival→*mu*]); as a result, (10) is ambiguous between 'It is not worthwhile for him to sell it (to someone)' and 'It is not worthwhile (for someone) to sell it to him'. To avoid ambiguity, either the infinitive clitics need to be placed locally (*Ne vredi* mu [*prodavati* ga] vs. *Ne vredi* [*prodavati* mu ga]). Better still, a different wording should be used, with a *da*-clause instead of the infinitive phrase in the first case, and the full form of the pronoun in the second (*Ne vredi* mu [*da* ga *prodaje*] vs. *Ne vredi* <ga> [njemu <ga> *prodavati*]).

2.2.3.2. CPC Resulting in a High Number of Clitics

(11) a. Ko **li** **bi** **<nas ga>** **se** usudio [lišiti
who INTERR COND us him REFL $dare_{PTCP}$ to.deprive
<nas ga>]?
'Who, I wonder, would dare to deprive us of it?'

b. (i) Bila **ga** **je** čast [poznavati **<ga>**].
be_{PTCP} him is honor to.know
'It was an honor to know him.'

(ii) Bila **mi** ??**<ga>** **je** čast [poznavati **<ga>**].
to me

c. Nepravedno/ Nepravda **<im** **ih>** **je** [nametati ?**<im ih>**].
unjust/ injustice to.them them is to.impose
'It is unjust/It is an injustice to impose them (e.g., 'these conditions') on them.'

The number of clitics seems to be less of a factor when V_{INF} is a verbal complement, as in (11a), which is fine both with the CPC of the infinitive clitics (resulting in a cluster of five clitics) and without it. With the V_{INF} as the subject, however, in situ placement of its clitics is preferable, at least in some cases, as in (11b–c).

2.2.3.3. CPC Resulting in Case and Case-Person Incompatibilities

(12) a. *[DAT+DAT]

Nemoguće **mi** *<**mu**> **je** [to <**mu**> objasniti].
impossible to.me to.him is that$_{(PRON)}$ to.explain

'It is impossible for me to explain that to him'.

b. *[ACC+ACC]

Sramota **me** *<**ga**> **je** [lagati <**ga**>].
shame me him is to.lie

'I am ashamed to lie to him.'

c. *[DAT+ACC(1/2P)]

Čast */?[**mi** <**vas**>] **je** [pozvati <**vas**>].
honor to.me you$_{PL}$ is to.invite

'It is an honor for me to invite you.'
(Title of a Yugoslav talk show from the 1980-ies)

In the situations illustrated in (12a–b), CPC is outright ungrammatical. In those exemplified in (12c), it may be marginally acceptable to some speakers.[15]

[15] It seems that in clitic pseudo-climbing environments two accusative clitics are allowed in Czech (Dotlačil 2007):

(i) *ale stále* nás ho$_i$ *nutila* [*jíst* __$_i$]
but always us$_{ACC}$ it$_{ACC}$ forced to.eat

'But she always forced us to eat it.'

This particular example would not sound that awful in Croatian, either: (ii) *ali uvek* nas ga je *terala jesti*.

Franks 2010: 134 gives a grammatical Polish example with the order ACC+DAT (while the normal order is DAT+ACC).

(ii) *Nauczyłeś* go mi$_i$ [*pomagać* __$_i$]
taught him$_{ACC}$ me$_{DAT}$ to help

'You taught him to help me.'

vs.

**Nauczyłeś* mi$_i$ go [*pomagać* __$_i$].

2.2.3.4. CPC Resulting in the Sequence *[*se je*]

(13) a. Ružno/ Sramota *<**se**> **je** [svađati <**se**>].
ugly/ shame REFL is$_{(V.COP)}$ to.quarrel
'It is ugly/It is a shame to quarrel.'

b. Ne prašta nikome ko <**mu** ~~**se**~~> **se** **je**
not forgives no.one who to.him ~~REFL~~ REFL ~~is~~$_{(V.AUX)}$
usudio [suprotstaviti <**mu se**>].
dare$_{PTCP}$ to.oppose
'He does not forgive anyone who dared to oppose him'.

In sentence (13a), the CPC of the reflexive marker results in the illegitimate and "irreparable" sequence *[se je$_{(V.COP)}$].[16] The sequence *[se je$_{(V.AUX)}$], however, can be repaired, because, unlike the copula, the auxiliary can be deleted; this is shown in (13b), which also illustrates the haplology of one of the two instances of the reflexive marker in the situation of non-local placement. (See examples (4b-i) and (4b-ii), as well.)[17]

Note finally that these may not be all the factors relevant for infinitive clitic placement. Thus, for instance, it seems that even word order within the infinitive phrase plays a role: Compared to *Buskeros ponovo poče* [*smejati* **se** *kao lud*] 'B. started laughing his head off again.' (example (3c) above), the variant *Buskeros ponovo poče* [*kao lud* **se** *smejati*] sounds more natural. However, I will not pursue this point further.

(An analogous example would not be correct in Serbian, unless the dative clitic is an Ethical Dative.) This indicates that some "normal" clitic cluster building rules may be transgressed when mixed clusters are constructed.

[16] While some Serbian speakers may tolerate the [**se je**] sequence (I thank a reviewer for pointing out to me some corpus data attesting to this), my feeling is that standard Serbian has a strong preference for avoiding it; perhaps the constraint is not that strict in the older language or in some contemporary regional variants. In Croatian, however, this clitic sequence is normal; cf. the following example, found on WWW: *Ružno* se je *osjećati bespomoćno/prejesti* (lit. ugly REFL is to.feel helpless/to.overeat) 'It is ugly to feel helpless/to overeat.'

[17] A *[*se je*] sequence can of course arise independently of CPC; e.g., with a reflexive verb in the past tense: *Setio* *[se je$_{V.AUX}$] (lit. [He] having.remembered REFL is) 'He remembered.' (⟹ *Setio* se). For more on this and other constraints on the co-occurrence of Serbian clitics, see Milićević 2007.

3. Linearization of the Infinitive Clitics in a Dependency-Based Approach to Syntax

First the theoretical framework is briefly presented (3.1), followed by sample infinitive clitic linearization rules (3.2).

3.1. Theoretical Framework

Within a Meaning-Text model of language L, a functional, stratification model making use of dependency-based linguistic representations, CPC is viewed as a particular case of linearization of clitic lexemes (the latter being itself a particular case of linearization of lexemes—full and clitic).[18]

The operation of linearization takes place in the transition between the Surface-Syntactic Representation (SSyntR), whose basic structure (SSyntS) is a linearly unordered dependency tree made up of surface lexemes subscripted with semantically full inflectional values and linked by binary surface-syntactic relations, and the Deep-Morphological Representation (DMorphR), its basic structure (DMorphS) being a fully ordered string of lexemes subscripted with all (semantically full and syntactically induced) inflectional values. Prosodic features are modeled, for a given clause, by the prosodic structures, superimposed on the basic representations of the two levels.

Cliticization, i.e., the marking of lexemes for the inflectional value CLIT (as opposed to FULL), is performed in the SSyntR $\Longrightarrow$ DMorphR transition. The actual production of clitic forms takes place at a later stage of the synthesis. (In other words, only the "sources" of clitics appear in the two structures.) In most cases, the communicatively unmarked auxiliaries and personal pronouns end up in the clitic form; i.e., CLIT is the default value of the inflectional category of tonicity.

The sources of clitics appear in the SSynt roles shown in Table 4 on the following page. As we can see, 2P clitics depend on different governors or are governors themselves; yet they are linearly positioned together with respect to a common host. That is, they do not behave like normal word forms, which are linearly positioned with respect to their SSynt-governors.

Since the SSynt-governor of a 2P clitic and its host are in principle distinct (although they may, of course, coincide), whether an infinitive clitic is placed locally or non-locally is not in any way reflective of the syntactic structure of the clause, which is the same in both cases. Thus with 2P clitics there is no genuine climbing and their local vs. non-local placement is strictly a linearization problem. In contrast, the SSynt-governor of an ad-verbal clitic of Romance type and its host always coincide, i.e., ad-verbal clitics are governed by the

[18] Recall that we are adopting the viewpoint of synthesis (speech production), rather than that of analysis (speech understanding).

Table 4. Surface-Syntactic roles of Serbian clitics

$L_{V.AUX(FIN)}$	top-node of the SSynt-tree, i.e., the head of the corresponding clause
$L_{PRON,\ CASE\ =\ GEN}$ <DAT, ACC>	object of $L_{V.LEX}$, L_N or L_{ADJ}
SE	adjunct of $L_{V.LEX}$
LI_{INTERR}	adjunct of L_V
LI_{EMPHAT}	adjunct of the lexeme L on which it bears semantically

verb and linearly positioned with respect to it. Therefore, to be non-locally positioned an ad-verbal clitic has to be made dependent on a different governor "higher" in the syntactic tree, and this can be metaphorically described as climbing. As a result, SSyntS of clauses featuring locally *vs*. non-locally positioned ad-verbal clitics are different.

The Serbian sentences in (14) have the same underlying SSyntS, shown in Figure 1; their DMorhpSs are given in Figure 2 on the following page.

(14) a. Tek **ga** **je** tada počeo [istinski uvažavati].
only him is then start$_{PTCP}$ truly to.appreciate
'Only then did he start truly to appreciate him.'

b. Tek tada **ga** **je** počeo [istinski uvažavati].

c. Tek **je** tada počeo [istinski **ga** uvažavati].

d. Tek tada **je** počeo [istinski **ga** uvažavati].

In (14a–b), the pronominal clitic *ga* is positioned non-locally (i.e., it pseudo-climbs); it forms a common cluster with the auxiliary clitic *je*, which is inserted into the first constituent of the DMorphS in (14a) and placed after this constituent in (14b).

In (14c–d), the infinitive clitic is positioned locally (i.e., does not pseudo-climb); it forms a cluster of its own, which sits in 2P of the infinitive phrase. The clause-level cluster consisting only of the auxiliary clitic is inserted into the first constituent of the DMorphS in (14c) and placed after this constituent in (14d).

Compare this situation with the genuine clitic climbing, as in the following Spanish example:

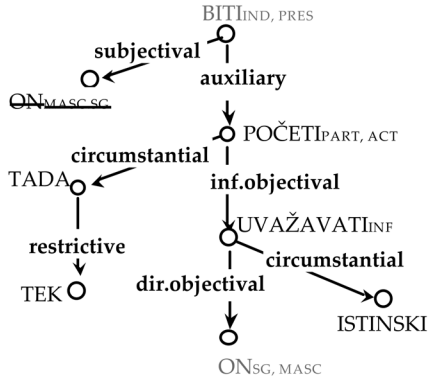

Figure 1. Common SSyntS of sentences (14a–d)

(15) a. Quiero poder encontrar$_{GOVERNOR\&HOST}$**==lo**.
'[I]want to.be.able to.meet him'.
b. Quiero poder$_{GOVERNOR\&HOST}$**==lo** encontrar.
c. **Lo==**quiero$_{GOVERNOR\&HOST}$ poder encontrar.

Here a change of SSynt-governor must take place for the clitic *lo* to be able to change the host; the sentences in (15) correspond to three distinct SSyntSs (and of course three distinct DMorhpSs).

To sum up, no syntactic rule is necessary to account for CPC of 2P clitics in Serbian and other languages featuring this type of clitics; it is taken care of by a subset of clitic linearization rules.

(14a) [TEK **ON**$_{\text{SG.MASC, IND, CL, 3SG}}$ **BITI**$_{\text{PRES, IND, CL, 3SG}}$ TADA] [POČEO$_{\text{PART, SG, MASC}}$] ||
[ISTINSKI UVAŽAVATI$_{\text{INF}}$]

(14b) [TEK TADA] **ON**$_{\text{SG.MASC, IND, CL, 3SG}}$ **BITI**$_{\text{PRES, IND, CL, 3SG}}$ [POČEO$_{\text{PART, SG, MASC}}$] ||
[ISTINSKI UVAŽAVATI$_{\text{INF}}$]

(14c) [TEK **BITI**$_{\text{PRES, IND, CL, 3SG}}$ TADA POČEO$_{\text{PART, SG, MASC}}$] ||
[ISTINSKI **ON**$_{\text{SG.MASC, IND, CL, 3SG}}$ UVAŽAVATI$_{\text{INF}}$]

(14d) [TEK TADA **BITI**$_{\text{PRES, IND, CL, 3SG}}$ POČEO$_{\text{PART, SG, MASC}}$] ||
[ISTINSKI **ON**$_{\text{SG.MASC, IND, CL, 3SG}}$ UVAŽAVATI$_{\text{INF}}$]

Figure 2. DMorphSs of sentences (14a–d)

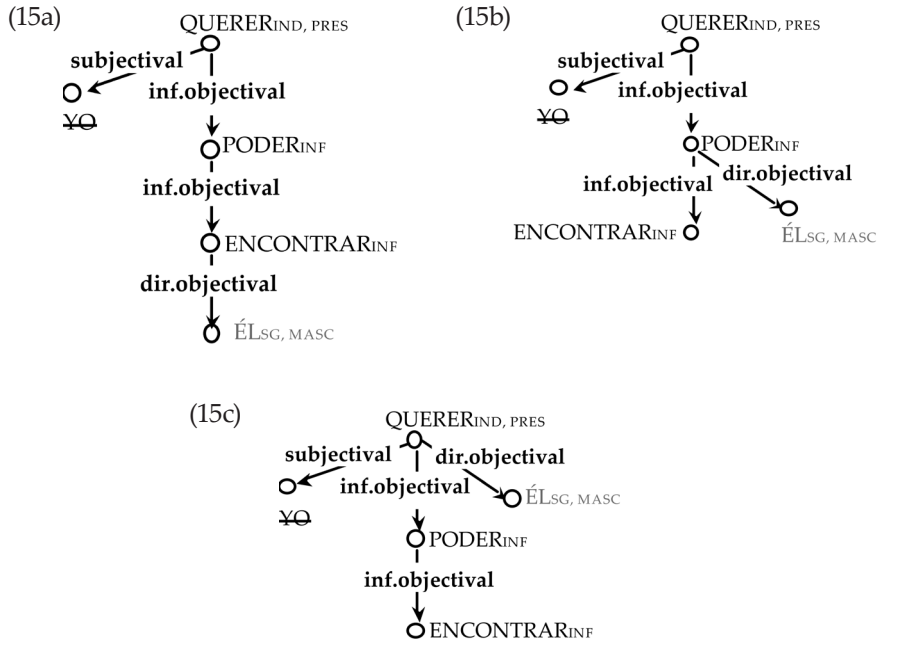

Figure 3. SSyntSs of sentences (15)

3.2. Sample Rules for the Linear Placement of the Infinitive Clitics

The following rules are needed to account for the linear placement of the infinitive clitics within the clause:

(1) Rules specifying the basic linear placement options for the infinitive clitics.
(2) Preference rules for the cases where both placement options apply.
(3) Filter rules, specifying the clitic sequences to be avoided when mixed clitic clusters are constructed.

Only the rules of the first type will be fully presented below. For the two other rule types sample rules will be sketched.

3.2.1. Rules Specifying the Basic Linear Placement Options for the Infinitive Clitics

The rules are of the form "if X, then Y". (This, again, is strikingly different from the linearization rules for normal words, which have in their left-

hand part a surface-syntactic subtree and in their right-hand side, a deep-morphological string.)

A RULES: OBLIGATORY NON-LOCAL PLACEMENT [= CPC obligatory]
[A.1] Non-local placement of the infinitive clitics is OBLIGATORY if:
a) V_{INF} = Lexical Part of MV in the future tense
OR
b) 1. The infinitive phrase is NOT clause initial and it is NOT heavy
AND
2. V_{INF} = Object of V | V is not lexically marked as not admitting CPC out of its inf. complement
[A.2] A single clitic cluster is constructed.
Examples: (2a) for A.1a; (1a), (3a–b) for A.1b; with V_{INF} lexically marked: (4)

B RULES: OPTIONAL LOCAL PLACEMENT [= CPC optional]
[B.1] Local placement of the infinitive clitics is OPTIONAL if:
a) 1. The infinitive phrase is NOT clause initial & it is heavy
2. V_{INF} = Object of V
OR
b) 1. The infinitive phrase is NOT clause initial
AND
2. V_{INF} = Subject of V or Object of N/ADJ
[B.2] Two separate clitic clusters are constructed.
Examples: (1b), (2b), (3c–e) for B.1a; (5–6) for B.1b.

C RULES: OBLIGATORY LOCAL PLACEMENT [= CPC blocked]
[C.1] Local placement of the infinitive clitics is OBLIGATORY if:
a) The infinitive phrase is clause initial
OR
b) V_{INF} = Copular Attribute
[C.2] a) A single clitic cluster is constructed if V_{INF} = Object of V.
b) Two separate clitic clusters are constructed if $V_{INF} \neq$ Object of V.
Examples: (1c) for C.2a; (1d) and (7) for C.2b.

All clusters are constructed according to the same cluster-building rules (Milićević 2009a: 264). Special filters apply to mixed clusters—to make sure that no ungrammatical or otherwise unacceptable clitic sequences occur; see 3.2.3.

All clusters are linearly placed (into the partial DMorphS of the clause (a fully ordered sequence of full-fledged clause elements) according to the same rules: 2P placement, skipping, insertion (Milićević 2009a: 267–69). Clause-initial infinitive phrases are subject to specific skipping rules; see 3.2.2.

3.2.2. Preference Rules

Only linguistic preferences proper, arising as a function of syntactic context, are considered here. Stylistic preferences (linked to the style/register) as well as sociolinguistic ones (dialectal/idiolectal) are not commented upon.

The heavier the infinitive phrase (and/or the rest of the clause), the more acceptable the local placement under rule B.1a.

In situations covered by rule B.1b, no inherent preferences exist. Local placement is preferred it if helps avoid non-desired consequences of non-local placement: ambiguity or cumbersome mixed clusters (with more than 2 clitics).

In situations covered by rule C.2b, the clause-initial infinitive phrase is preferably skipped, especially if it contains non-clitic elements, i.e., the clause-level cluster preferably does not immediately follow the infinitive phrase. Instead, a prosodic break is made and the clitic cluster put in the 2P of the rest of the clause; a resumptive *to*$_{PRON}$ 'that' is sometimes inserted; see (16a). However, immediate post-position of the clause-level clitic cluster consisting only of the copula is tolerated in some contexts; see (16b–c).

(16) a. [Tražiti **ga** u mraku] || teško **je** <|| to **je** teško>
to.seek him in darkness difficult is that is difficult
'Looking for him in the dark is difficult <, that is difficult>.'

b. [Tražiti **ga**] **je** teško.
to.seek him is difficult
'Looking for him is difficult.'

c. [Približiti **se** stražaru] **je** zabranjeno. (Popović 1997: 351)
to.approach REFL to guard is forbidden
'Approaching the guard is forbidden.'

3.2.3. Filter Rules

These rules identify the ungrammatical clitic sequences in mixed clusters: identical cases (ex. (12a–b) above); person-case incompatibility (ex. (12c)); *[se je$_{V.COP/AUX}$] sequence (ex. (13a)); *[se se] sequence (ex. (13b)).

Let us show how these rules work on sentences from (14) above, whose partial DMorphSs is shown in Figure 4 on the following page.

The infinitive phrase *istinski uvažavati*$_{INF}$ is not clause-initial, is heavy, and may be preceded by a prosodic break. This makes it possible to apply rules B.1a, allowing for the optional local placement of the infinitive clitic.

The non-local placement option is chosen in (14a–b), and the local placement option in (14c–d). In the first case, the mixed clitic cluster is constructed.

Full-fledged sentence elements

[TEK TADA]$_{+insert}$ [POČEO$_{PART, SG, MASC}$] || [ISTINSKI UVAŽAVATI$_{INF}$]$_{+heavy}$

Clitic cluster(s)

For non-local placement

[ON$_{SG.MASC, IND, CL, 3SG}$ BITI$_{PRES, IND, CL, 3SG}$]$_{CL.CLUSTER-1, MIXED}$

For local placement

[ON$_{SG.MASC, IND, CL, 3SG}$]$_{CL.CLUSTER-2, INF}$
[BITI$_{PRES, IND, CL, 3SG}$]$_{CL.CLUSTER-3, CLAUSAL}$

Figure 4. Partial DMorphSs for the construction of sentences (14)

It is placed after the first constituent of the clause in (14a) and inserted into that constituent in (14b);. This operation is taken care of by the rules handling all clusters (whether mixed or not). In the second case, two clitic clusters are constructed, the infinitive phrase and the clausal-level one, and each is placed in the 2P of the corresponding domain.

There are no constraints on and no obvious linguistic context-induced preferences for any of the linear arrangements illustrated in (14).

4. Conclusion

The paper examined, from a dependency-syntax vantage point, the linear placement options available for clitic dependents of infinitive complements in Serbian. It focused in particular on the non-local placement option, i.e., linear positioning of the infinitive clitics outside of the infinitive phrase, commonly referred to as clitic climbing. It was shown that, unlike an ad-verbal clitic, a 2P clitic does not need to change its syntactic governor in order to be positioned non-locally, i.e., that it does not undergo genuine clitic climbing. This is because the syntactic governor of a 2P clitic (in the linearly non-ordered dependency structure of the clause) and its host (in the fully ordered morphological structure) are in principle distinct, while in the case of an ad-verbal clitic they necessarily concide. Therefore, a non-locally positioned 2P clitic can only be said to display climbing effects, or to pseudo-climb.

The paper offered a mixed account of the linear placement of Serbian infinitive clitics, acknowledging the importance of both syntactic and prosodic factors. These findings are in line with the conclusions in Milićević (2009a) describing Serbian 2P clitic placement in general, although they point to a somewhat more important role of prosody than previously assumed.

Acknowledgments

My heartfelt thanks go to Wayles Browne and Igor Mel'čuk for their comments on a prefinal version of this paper, as well as to my BCS speaking friends for their help with judging the acceptability of some difficult examples. Thanks are also due to two reviewers, whose remarks helped me improve some aspects of the paper.

References

Aljović, Nadira. (2006) "Are there optional rules in syntax? Evidence from clitic climbing in Bosnian/Croatian/Serbian". Paper presented at the FDSL 6.5 Conference, Nova Gorica, Slovenia, 1–3 December 2006.

Bošković, Željko. (2004) "Clitic placement in South Slavic". Steven Franks, Uwe Junghanns, and Paul Law, eds. *Journal of Slavic Linguistics* 12(1–2): 37–90.

————. (2015) "On second position clitics cross-linguistically." Unpublished ms., University of Connecticut.

Browne, Wayles. (1975) "Serbo-Croatian clitics for English-speaking learners". *Kontrastivna naliza engleskog i hrvatskog ili srpskog jezika*, Zagreb: Institut za lingvistiku Filozofskog fakulteta, 105–134. [Reprinted in *Journal of Slavic linguistics* 12(1–2): 249–83.]

————. (1987) "Classification of subordinate clauses in a grammar of Serbo-Croatian for foreign users". *The Zagreb English-Serbo-Croatian project*. *Contrastive analysis of English and Serbo-Croatian* III. Zagreb: Institute of linguistics, 165–91.

Caink, Andrew. (2004) "Semi-lexical heads and clitic climbing". *Journal of Slavic linguistics* 12(1–2): 95–139.

Ćavar, Damir and Chris Wilder. (1999) "Clitics third in Croatian". Henk Van Riemsdijk, ed. *Clitics in the languages of Europe*. Berlin: Mouton de Gruyter, 429–67

Čamdžić, Amela and Richard Hudson. (2002) "Serbo-Croat-Bosnian clitics and word grammar". *University College London working papers in linguistics*m 14: 321–53.

Dobnik, Simon and Robin Cooper. (2016) "Restructuring restructuring". Franc Marušič and Rok Žaucer, eds. *Formal studies in Slovenian syntax*. *In honor of Janez Orešnik*. Amsterdam/Philadelphia: John Benjamins, 69–90.

Dotlačil, Jakub. (2007) "Why clitics cannot climb out of CPs: A discourse approach". *Formal approaches to Slavic linguistics: The Toronto meeting, 2006*. Ann Arbor, MI: Michigan Slavic Publications, 76–93.

Franks, Steven. (2010) "Clitics in Slavic". *Glossos* 10: 1–157.

Franks, Steven and Tracy Holloway King. (2000) *A handbook of Slavic clitics*. Oxford: Oxford University Press.

Golden, Marija. (2003) "Clitic placement and clitic climbing in Slovenian". *Language typology and universals*, 56(3): 208–33.

Halpern, Aaron. (1995) *On the placement and morphology of clitics*. Stanford, CA: CSLI Publications.

Hansen, Björn, Zrinka Kolaković, and Edyta Jurkiewicz-Rohrbacher. (2018) "Clitic climbing and stacked infinitives in Bosnian, Croatian, and Serbian. A corpus-driven study". Eric Fuß, Marek Konopka, Beata Trawiński, and Urlich Waßner, eds. *Grammar and corpora 2016*. Heidelberg: Heidelberg University Publishing, 259–68.

Jurkiewicz-Rohrbacher, Edyta, Zrinka Kolaković, and Björn Hansen. (2017a) "Web corpora—the best possible solution for tracking rare phenomena in underresourced languages: Clitics in Bosnian, Croatian, and Serbian". Piotr Bański et al., eds. *Proceedings of the workshop on challenges in the management of large corpora and big data and natural language processing*, 49–55.

————. (2017b) "Clitic climbing, finiteness, and the raising-control distinction. A corpus-based study". *Journal of Linguistics*, 68(2): 179–90.

Krapova, Iliyana and Giuseppina Turano. (2015) "On clitic climbing in the Balkan languages from a synchronic perspective". Antonio D'Alessandri and Francesco Guida, eds. *L'Europa et il suo Sud-est. Persorsi de ricerca. Contributi italiani all'XI congresso internazionale dell'Association internationale d'études du Sud-est européen*, 135–53.

Marušič, Franc. (2008) "Slovenian clitics have no unique syntactic position". *Formal approaches to Slavic linguistics*, 16, 266–281.

Mel'čuk, Igor. (1988) *Dependency syntax. Theory and practice*. Albany, N.Y.: State University of New York Press.

————. (2013) *Semantics. From meaning to text*, vol. 2. Amsterdam/Philadelphia: John Benjamins.

Milićević, Jasmina. (1999) "Pronominal and verbal clitics in Serbian: A morphological description". *Wiener Slawistischer Almanach* 43: 231–56.

————. (2007) "Co-occurrence of Serbian second-position clitics: Syntactic and morphonological constraints". Nabil Nathout and Fabio Montermini, eds. *Morphologie à Toulouse: Actes du colloque international de morphologie 4e Décembrettes*. München: Lincom Europa, 99–121.

————. (2009a). "Linear placement of Serbian clitics in a syntactic dependency framework". Alain Polguère and Igor Mel'čuk, eds. *Dependency in linguistic description*. Amsterdam: John Benjamins, 235–76.

————. (2009b). "Serbian auxiliary verbs—syntactic heads or dependents?" Wladyslaw Cichocki, ed. *Proceedings of the 31st annual conference of the Atlantic provinces linguistic association. Fredericton, 2–3 November 2007. PAMAPLA* 31: 43–53.

Mišeska-Tomić, Olga. (2006) "Infinitives and subjunctives". Olga Mišeska-Tomić, ed. *Balkan Sprachbund morpho-syntactic features*. Dordrecht: Springer; 413–510.

Mrazovac, Pavica and Zorica Vukadinović. (2009) *Gramatika srpskog jezika za strance*. Sremski Karlovci/Novi Sad: Izdavačka knjižarnica Zorana Stojadonovića.

Monachesi, Paola. (1998) "Italian restructuring verbs: A lexical analysis". Erhard Hinrichs, Andreas Kathol, and Tsuneko Nakazawa, eds. *Complex predicates in non-derivational syntax*. New York, NY: Academic Press, 313–68.

Nišida, Chiyo. (2012) "Restruturing and clitic climbing in Romance: A categorial grammar analysis". Artemis Alexiadou, Tibor Kiss, and Gereon Müller, eds. *Local modeling of non-local dependencies in syntax*. Berlin: Walter de Gruyter, 371–400.

Nušić, Branislav. (1924) *Autobiografija*. Available: www.lektire.me/prepricano/branislav-nusic-autobiografija_373.

Piper, Predrag, Ivana Antonić, Vladislava Ružić, Sreto Tanasić, Ljudmila Popović, and Branko Tošović. (2005) *Sintaksa savremenoga srpskog jezika: prosta rečenica*. Beograd: Beogradska knjiga/Novi Sad: Matica srpska.

Popović, Ljubomir. (1997) *Red reči u rečenici*. Beograd: Društvo za srpski jezik i književnost Srbije.

Progovac, Ljiljana. (1993) "Locality and subjunctive-like complements in Serbo-Croatian". *Journal of Slavic Linguistics*, 1, 116–144.

————. (2005) *Syntax of Serbian: Clausal architecture*. Bloomington, IN: Slavica Publishers.

Rizzi, Luigi. (1982) *Issues in Italian syntax*. Dordrecht: Foris.

Simić, Radoje. (1999) *Osnovi sintakse srpskog jezika*. *Vol. 1: Teorija iskaznih formi*. Nikšić: Jasen.

Spencer, Andrew. (1991) *Morphological theory. An introduction to word structure in Generative Grammar*. Oxford: Blackwell.

Stjepanović, Sandra. (1998) "On the placement of Serbo-Croatian clitics: Evidence from VP ellipsis". *Linguistic inquiry* 29: 527–37.

————. (2004) "Clitic climbing and restructuring with 'finite clause' and infinitive complements". Steven Franks, Uwe Junghanns, and Paul Law, eds. *Journal of Slavic Linguistics* 12(1–2): 177–213.

Zimmerling, Anton. (2011) "Scrambling types in the Slavic languages". *Computational linguistics and intellectual technologies* 10(17): 750–66.

Zimmerling, Anton and Peter Kosta. (2013) "Slavic clitics. A typology". *STUF-Language typology and universals* 66(2): 178–214.

Department of French
Dalhousie University
Room 1114 Marion McCain Arts and Social Sciences Building
6135 University Avenue
Halifax (Nova Scotia) B3H 4R2
jmilicev@dal.ca

Collocational Competence among Polish Students of Croatian as a Foreign Language

Antonia Ordulj and Nikolina Sokolić

Abstract: Collocations in the Croatian language have been described in detail (Blagus Bartolec 2014). However, research on collocational competence in Croatian as a first or foreign language is still sparse. Previously, only the factors that influence the collocational competence of Croatian native speakers (NSs) were investigated, showing that NSs have the best knowledge of collocations with high frequency and associative strength (Ordulj and Cvikić 2017). The most extensive research on collocational acquisition in Croatian as a foreign language (CFL) in heterogeneous groups was done by Ordulj (2017). This research showed that participants with lower proficiency had very poor collocational competence, while participants with higher proficiency showed an equal knowledge of noun collocations in the nominative case and in oblique cases. This paper aims to deepen previous findings and to examine productive knowledge of noun collocations based on their frequency, associative strength, morphological features, and the proficiency level of Polish students of Croatian. Collocations used in this research were collected from essays written by students of Croatian at B1 and B2 CEFR (2005) proficiency levels. The hrWaC corpus was used to count the frequency of collocations, and the associative strength of collocational constituents was assessed by native speakers of Croatian. The respondent sample comprised 27 students of Croatian in Krakow, Poland, who were divided into two proficiency level groups: lower (2nd and 3rd year of studies) and higher (4th and 5th year of studies). The influence of morphological features on collocational knowledge was tested with two fill-in-the-blank tasks containing collocations in the nominative and oblique cases. The analysis of productive tasks showed that morphological features do not influence collocational knowledge at the lower or higher proficiency level. In both cases participants produced the best results in tasks with collocations of high frequency.

1. Introduction

Collocational competence, as part of lexical competence, is one of the most important and challenging aims in Second Language Acquisition (SLA). It was Hill (1999) who introduced the notion of collocational competence into SLA in the late 1990s, thus shifting the focus from teaching individual lexical items to

developing collocational competence. Many applied linguists argue that collocations remain a neglected area of lexical competence development (El-Dakhs 2015; Martyńska 2004; Koya 2003; Gitsaki 1999), and unsurprisingly, the most extensive research has been done on collocations in English as a Second Language (ESL). In vocabulary acquisition, emphasis has mainly been placed on acquiring individual lexical units, although it is estimated that about 70% of everything we hear, say, read, or write falls into the category of collocations (Hill 2000: 53). The importance of collocations in SLA has been deservedly recognized by applied linguists (Nation 2001; Hill 2000; Gitsaki 1999; Lewis 1997). For instance, Nation (2001: 517) claims that "language knowledge is collocational knowledge" and that "all fluent and appropriate language requires collocational knowledge". Consequently, it goes without saying that collocational usage in a foreign language contributes to fluency and more natural communication (Shin 2007; Borić 2004). Collocations also narrow the meaning of individual lexical units, which enhances comprehension and production in a foreign language (Schmitt 2004). Additionally, the use of structurally and semantically different types of collocations contributes to a more interesting style in oral and written production. As was just mentioned, collocations are often considered a neglected part of vocabulary acquisition. On the one hand, students are likely to learn and use individual lexical units without paying attention to the environment in which they naturally co-occur, thus contributing to insufficient development of reading skills. On the other hand, teachers of foreign languages often emphasize grammar, especially in morphologically rich languages, which is not beneficial for students' development of lexical and collocational competence. Another issue in collocational research is the lack of a clear and unambiguous definition of a collocation and the resulting inconsistencies in terminology, as well as different approaches to the phenomenon, which makes collocations difficult to understand even for teachers.

This study was motivated by the general lack of research on the collocational competence of non-native speakers of Croatian. Since research on collocational competence in CFL is still rather sparse (Ordulj 2016, 2017; Burić and Lasić 2012; Petrović 2007), this study aims to shed light on the productive knowledge of Croatian noun collocations among Polish students of CFL, considering their collocational frequency, associative strength, and morphological features. In a broader context, this research also provides insight into the acquisition of collocations in cases where students' first language is morphologically complex and typologically similar to the target language. The results and instruments used in this study could be useful in designing teaching materials and research instruments with different collocational types and for developing new approaches to teaching CFL, specifically for developing collocational competence in heterogeneous and homogeneous groups of CFL learners.

In the following sections, the notion of collocations in relation to other lexical combinations (idioms and free lexical combinations) in Croatian is presented, followed by a review of relevant literature. In the second part of the paper, the research questions, hypotheses, instruments, and methodology of the present study are described in detail, followed by the interpretation of the findings on collocation knowledge among Polish students of CFL.

2. Background

2.1. Theoretical Background

Previous studies on collocations in ESL demonstrated a pluralism of definitions and approaches to this lexical phenomenon. The same can be said for Croatian linguistics, which has shown a growing interest in collocations in recent years. The majority of studies are concerned with the lexicographic presentation of collocations, most likely due to the fact that Croatian still lacks a collocational dictionary. Generally speaking, research on collocations in Croatian can be divided into several strands that deal with collocational structure (Petrović 2007, 2008b), theoretical issues (Stojić 2012; Ivir 1992–93), contrastive and semantic approaches to collocations (Borić 1998), lexicographic processing of collocations (Turk 2010; Borić 2002; Pritchard 1998; Petrović 2008a; Duplančić Rogošić 2007), and collocations in language for specific purposes (Štefić, Mravak-Stipetić, and Borić 2010; Miščin 2012). Since there is still no clear and unambiguous definition of collocations, both the Croatian and the foreign literature list the features as formal criteria for identifying collocations. Features such as restrictions on the combination of constituents and greater probability or predictability of co-occurrence and recurrence (Borić 1998, 2002; Pritchard 1998; Benson 1985) cannot be considered the only criteria for differentiating collocations from other lexical units, because these features are inherent in all lexical and grammatical word combinations in a communicative context. Furthermore, Croatian linguists often ignore the semantic (analyses of semantic relations between collocational constituents) and formal features (such as parts of speech, morphological, syntactic, and word formation features) that are most relevant for defining collocations in Croatian. All the above points call for observing collocations in relation to other lexical combinations, especially idioms and free lexical combinations, which seem to represent the largest stumbling block for linguists in defining collocations without vagueness.

In Croatian linguistics, idioms are seen as expressions of secondary nomination, which indicate personal and emotive attitudes and opinions (Turk 2000), and do not arise in spoken language (i.e., by spontaneous word arrangement) but are rather incorporated in discourse as fixed expressions. The nature of collocations is seen as rather arbitrary, although collocational constituents

occur together in a fixed order. On the other hand, free lexical combinations are defined as sequences of words that are characterized by many possibilities of combining and exchanging elements, and therefore, have a lower possibility of co-occurrence (Petrović 2007; Borić 2002; Pritchard 1998). In this research, however, collocations are differentiated from free lexical combinations by taking into account the semantic criterion and freedom of combination. Those lexical items whose elements can be combined with other units without limits are considered free lexical combinations. Collocations are lexical combinations that arise from the syntagmatization process, where assimilation or change in meaning of individual elements creates new collocational meaning.

The most comprehensive definition of collocations in Croatian linguistics, which is also accepted in this paper, was given by Blagus Bartolec (2014: 80), who argues that a collocation is a "special lexical combination at the syntagmatic level based on semantic connectedness of individual lexical units, which specifies their meanings". She further emphasizes the communicative function of collocations, since they are very frequent among native and non-native speakers, which differentiates them from other types of lexical combinations. When it comes to the semantic features of collocations in Croatian, it should be mentioned that they do not have the status of lexical units, given that they arise from the process of syntagmatization. According to Blagus Bartolec (2014: 85), a crucial characteristic of collocations is the semantic potential of their constituents. In other words, constituents in the collocational relationship come with their own independent meanings, and through interaction they create collocational meanings that point to extralinguistic reality according to the communicative needs of the speaker. This functional approach to collocations is also accepted in this paper. However, since collocations in the context of CFL are sparsely studied, this research focuses only on collocations having nouns as bases and adjectives as their collocators.

2.2. Collocational Competence in Foreign Language

One of the major foci of studies on collocational competence in ESL is to examine receptive and productive knowledge of collocations. For instance, Begagić (2014) used three fill-in-the-blank productive tests and appropriateness-of-judgment receptive tests to investigate knowledge of verb-noun, adjective-noun, and verb-adverb collocations among 40 students majoring in English Language and Literature in Zenica, Bosnia and Herzegovina, who are native speakers of Bosnian, Croatian, and Serbian. Begagić (2014) found that students have generally inadequate collocational knowledge, with receptive knowledge outranking productive, and that the most problematic collocation types were verb-adverb and adjective-noun. According to Begagić (2014), students' lack of awareness of the importance of collocations (since they are generally focused on learning grammar rules), as well as negative transfer from their first lan-

guage could be the possible factors causing such results. Poor collocational knowledge of Polish-speaking intermediate English learners was shown by Martyńska (2004). In her study, participants demonstrated satisfactory knowledge of individual lexical units, but they knew only 55% of all collocational types, excelling in verb-noun and adjective-noun collocations.

In a study conducted on Spanish-speaking university students of English, Jaén (2007) confirmed students' better receptive knowledge of adjective-noun collocation types. According to Jaén, one of the most important aspects of testing collocational competence is the selection of testing items. She argues for using corpora as a source, relying on frequency of collocational constituents. In her opinion, corpora are the most reliable sources for the selection of items because they contain "authentic and representative samples of the language" (Jaén 2007: 135). Regarding other criteria for testing, it is important that collocations are used in different "text types and contexts, [and that] they are semantically transparent and restricted in their commutability" (Jaén 2007: 136). The importance of corpora as the most reliable and useful sources for selecting collocations and counting their frequency was confirmed by Siyanova-Chanturia and Spina (2015: 551). However, apart from the usage of corpora, they argued that teachers' intuition may represent a very important factor in estimating collocational frequency. Their stance was corroborated in a comparative study of intuitive judgments of adjective-noun collocation frequency among native speakers and learners of Italian (Siyanova-Chanturia and Spina 2015). In general, their results revealed that L2 learners of advanced proficiency in Italian were equally successful in estimating collocations of high and low frequency as native speakers. These findings suggest that both native speakers and L2 learners are successful in estimating lexical items at the "far ends of the frequency continuum" (Siyanova-Chanturia and Spina 2015: 552). This follows usage-based theories, according to which acquisition of lexical items is determined by frequency of (co-)occurrence in language input (Ellis 2012). Using items of high frequency creates stronger mental representations, so the most problematic items remain those in the middle of the frequency continuum because "they are less salient and less striking than highly frequent or infrequent items" (Siyanova-Chanturia and Spina 2015: 552).

Although large native corpora can be representative sources of collocations, there still remains the question of how reliable frequencies collected by large native corpora can be in the collocational acquisition of L2 learners. It should be taken into account that non-native speakers are mostly exposed to classroom language instructions and that their usage of linguistic structures and patterns differs significantly from that of native speakers, who are immersed in a natural language environment. Due to their limited exposure to language, and especially to collocations, it often occurs that infrequent expressions for NSs are much more frequently used by NNSs and vice versa. For that reason, non-native speakers often underuse or overuse some collocations

that they are familiar with, regardless of frequency in native corpora or in a natural language environment. To address this question and other factors influencing productive knowledge of English collocations sampled from the COCA, including frequency, mutual information (MI) and amount of language study and instruction, González Fernández and Schmitt (2015) tested Spanish speakers of EFL using a productive form recall format. The following findings emerged:

Firstly, corpus frequency has a moderate influence on collocational knowledge, but higher than MI, although it must be noted that "frequency based on general English native speaker corpora may not be the best way of sequencing collocations in instruction, as it may not reflect actual learner exposure very well." (González Fernández and Schmitt 2015: 114)

Secondly, consistent reading, watching English TV/films and using social networking sites are very useful activities for developing collocational competence in EFL.

Thirdly, the amount of language study and instruction moderately influences collocational knowledge.

The importance of choosing which collocations to teach was also pointed out by Hill (2000). Taking into consideration the criterion of collocational strength, Hill (2000: 63–64) differentiates strong, weak, and medium-strength collocations. Strong collocations encompass those collocational constituents that can be replaced with only a limited number of lexical items (e.g., *rancid butter*). Weak collocations are characterized by a somewhat weaker predictability and for that reason sometimes they can easily be confused with free lexical combinations (e.g., *white wine*). The most important collocations, according to Hill (2000: 63–64), are medium-strength collocations, because they are in the middle of the collocation spectrum and because learners usually know individual lexical items but not the collocations made of these frequent words.

Taking into consideration the criterion of restrictedness, Nesselhauf (2003) analyzed verb-noun collocations in essays written by German learners of English. The results showed that learners struggled most with lexical combinations with a medium degree of restrictedness, where the verb takes a wider range of nouns (e.g., *exert influence, control, pressure, authority, power, attraction* …), which means that these combinations are "more creatively combined by learners" (Nesselhauf 2003: 233). The best-learned combinations were those with a high level of restrictedness (e.g., *pay attention*), because they are "more often acquired and produced as wholes" (Nesselhauf 2003: 233).

It is believed that the development of general vocabulary increases productive and receptive collocational knowledge. This has been confirmed in research on English (Koya 2003; Gitsaki 1999) and Italian as L2 (Bonci 2002). Furthermore, collocational usage can be a clear indicator for differentiation of lower and higher proficiency levels of non-native speakers. Students at

higher ESL proficiency levels use more complex types of collocational structures than students at lower proficiency levels, who mostly rely on frequent collocations consisting of nouns, adjectives, and verbs (Gitsaki 1999). This could be due to exposure to collocations, differences in the first language, the structural and semantic complexity of collocational types but also due to educational context, that is, the exposure to different collocational types in the learning process (Begagić 2104; Gitsaki 1999). Previous research has shown that receptive knowledge of collocations comes before productive, but collocational competence in ESL is generally rather insufficiently developed and neglected (Begagić 2014; Jaén 2007; Martyńska 2004). Additionally, it is evident that choosing examples of collocations appropriate for research based on certain criteria (e.g., frequency, collocational strength, restrictedness) is crucial. When it comes to the most appropriate source of collocations, Jaén (2007) argues for corpora, whereas Siyanova-Chanturia and Spina (2015) emphasize teachers'/native speakers' intuition.

2.3. Collocational Competence in Croatian as a Foreign Language

Research on collocational competence in Croatian as a foreign language (CFL) is still sporadic. The few studies that investigate the phenomenon focus on different types of exercises for developing collocational competence (Petrović 2007), or on different types of collocations in textbooks for CFL beginners (Burić and Lasić 2012), noting that the most frequent collocations in CFL are adjective-noun collocations. A recent detailed study on factors influencing productive knowledge of adjective-noun collocations in CFL was done by Ordulj (2017). It considered the frequency of collocations, their associative strength (strong and weak collocations), morphological features (collocations in the nominative case and morphologically marked collocations in oblique cases), as well proficiency (B1 and B2 CEFR levels). It should be emphasized that the sample structure, comprised of 70 students of B1, B1+, B2, and B2+ proficiency levels, was very heterogeneous regarding gender, language (pre) knowledge, and first language. Based on her results, Ordulj (2017) formulates the following general conclusions:

Firstly, participants at B1 proficiency level show better production of frequent collocations in the nominative than in oblique cases. More precisely, they show very weak knowledge of morphologically marked collocations in oblique cases, regardless of their frequency and associative strength. Therefore, the most important factors influencing the use of collocations among B1 proficiency participants are morphological features and frequency.

Secondly, participants at B2 proficiency level show equal knowledge of noun collocations in the nominative and oblique cases, which means that morphological features do not significantly affect their collocational knowledge. Participants had the best results in tasks with strong collocations of higher

frequency, which corroborates frequency and associative strength as the most important factors for better collocational knowledge among B2 proficiency participants. This is in line with a study of the collocational competence of native speakers of Croatian done by Ordulj and Cvikić (2017).

Thirdly, response-accuracy analysis of tasks with collocations in a wider context for B1 and B2 CFL proficiency levels shows that context does not have a significant influence on collocational usage. Participants at the B2 proficiency level performed better with strong collocations of higher frequency and weak collocations of lower frequency, but surprisingly both B1 and B2 CFL learners knew weak collocations of low frequency better than weak collocations of higher frequency.

For the purpose of this research, it should be mentioned that Polish and Croatian are typologically similar languages, with resemblances reflected on all linguistic levels. Both of these Slavic languages have grammatical person, number, gender, case, aspect, and so on. Despite the similarities between the two languages, Polish-speaking learners of CFL still have difficulties at all levels of language proficiency, as reported by Aleksovski 2014, Kordić and Vidović Bolt 2013, and Podboj 2013. For instance, there are difficulties in acquiring Croatian noun and adjective agreement, since many nouns in Polish have similar or even the same form as their Croatian counterparts, very often with the same meaning but different grammatical gender (Aleksovski 2014). Furthermore, problems relating to meaning and usage of words in typologically related languages are frequent, since their contents are usually only partially the same, and many cases of false friends additionally complicate the issue (see Sokolić and Vidović Bolt 2012 for a more detailed review).

3. Research Problems and Hypotheses

In this study, productive knowledge of Croatian adjective-noun collocations among high and low proficiency level Polish students of CFL is investigated, considering word frequency, associative strength, and morphological features. There are several reasons why the acquisition of collocations in CFL is examined considering the aforementioned factors. Firstly, as can be seen from previously mentioned studies (Siyanova-Chanturia and Spina 2015; Ellis 2012; Jaén 2007; Nation 2001), frequency is one of the most important factors in selecting collocations for research, since it directly affects the usage of lexical units. As Ellis (2002: 144) points out: "Frequency is a key determinant of acquisition because "rules" of language, at all levels of analysis (from phonology, through syntax, to discourse), are structural regularities that emerge from learners' lifetime analysis of the distributional characteristics of the language input". Secondly, collocational strength is a key factor in choosing examples for classroom practice and evaluation, as was noted by Hill (2000). Since previous research on collocational competence of native speakers (Ordulj and

Cvikić 2017) and a heterogeneous group of non-native learners of Croatian (Ordulj 2017) had confirmed that the use of noun collocations is under the influence of frequency and associative strength, it is reasonable to assume that these factors will influence the performance of Polish CFL learners as well. Learning collocations is based on the associative learning of sequences because "objects once experienced together tend to become associated in the imagination" (James 1950: 363) and chunking is the mechanism that supports language reception and production. As Pawley and Syder (1983: 192) point out, "In the store of familiar collocations there are expressions for a wide range of familiar concepts and speech acts, and the speaker is able to retrieve these as wholes or as automatic chains from long-term memory". On the other hand, acquisition of lexical items in CFL is affected by morphological diversity, that is, noun and adjective declension, phonetic assimilation, and morphological homonyms and homographs (Cvikić and Bošnjak 2004). Since previous research has already confirmed that morphological features influence the acquisition of collocations at B1 CFL proficiency level (Ordulj 2017) and that Polish CFL learners can have difficulties in acquiring Croatian noun and adjective agreement (Aleksovski 2014; Kordić and Vidović Bolt 2013), it can be assumed that morphology could be one of the factors influencing the use of noun collocations among Polish CFL learners.

Basing this research on results previously presented by Ordulj (2017) with non-native speakers of CFL and taking into account the aforementioned factors, the main goal of this study was to examine knowledge of noun collocations among Polish CFL learners of higher and lower proficiency levels based on collocational frequency, associative strength, and morphological features.

The following research questions and hypotheses will be addressed:

Q1: Do morphological features influence productive knowledge of noun collocations among participants of lower and higher CFL proficiency levels?

H1a: At the lower proficiency level, morphological features will affect productive knowledge of noun collocations. Participants will perform better on tasks with morphologically unmarked collocations in the nominative case compared to morphologically marked collocations in oblique cases.

H1b: At the higher proficiency level, morphological features will not affect productive knowledge of noun collocations. Participants will perform equally well on tasks with collocations in the nominative and oblique cases.

Q2: Do frequency and associative strength influence productive knowledge of noun collocations among participants of lower and higher CFL proficiency levels?

H2: Participants with higher CFL proficiency level will have more correct answers than participants with lower CFL proficiency level. Furthermore, at the lower proficiency level, frequency of collocations will be crucial, causing lower proficiency participants to have more answers that are correct in tasks with high rather than low frequency, whereas the associative strength will not make a significant difference. On the other hand, at the higher proficiency level, frequency and associative strength will affect collocational knowledge, with participants with higher proficiency excelling in tasks with collocations of high frequency and strong associative strength, compared to collocations of high frequency and weak associative strength, and low frequency and strong/weak associative strength.

Q3: Do frequency and associative strength influence knowledge of noun collocations in the wider context taking into account CFL proficiency levels?

H3: Participants with higher proficiency will have more correct answers than participants of lower proficiency. The frequency and associative strength of collocational constituents will have an effect on productive knowledge of collocations at both proficiency levels. Participants will perform better on tasks with collocations of high frequency and strong associative strength, compared to all other types of collocations. Still, answer accuracy will be higher for collocations of lower frequency and weak associative strength compared to collocations of higher frequency and weak associative strength.

4. Methodology

4.1. Participants

Participants in this study were university students of Croatian in Krakow, Poland. The sample consisted of 27 learners (10 male and 17 female participants) with an average age of 23 years (SD = 1.68). Their years of study were distributed as follows: five students were in the second year (A2 level), nine students were in the third year (B1 level), five students were in the fourth year (B2 level), and eight students were in the fifth year (C1 level) of their studies. It should be emphasized that their first language is Polish, which means that the

sample was homogeneous in that aspect. Since the sample was rather small, for the purposes of testing the hypotheses it was divided into two proficiency level groups: lower (2nd and 3rd year) and higher (4th and 5th year).

4.2. Materials: Criteria for Collecting and Selecting Collocations in the Present Study

As already mentioned, adjective-noun collocations chosen as testing items were selected according to the following criteria: semantic and formal features, frequency, and associative strength. The reasons for selecting collocations according to these criteria are described in section 3. Semantic and formal features for defining collocations in Croatian according to Blagus Bartolec (2014) are described in section 2.1. The division into base (noun) and collocator (adjective) allowed for a further corpus analysis of selected collocations.

Collocations were collected from 300 essays on different topics written by learners of CFL at B1 and B2 proficiency levels. Studying collocations and collocational competence in Croatian is still rather difficult because there is no collocational dictionary for either native or nonnative speakers and because other sources of collocations are limited. The first Learner Corpus of Croatian as a Second and Foreign Language—CROLTEC (Mikelić Preradović, Berać and Boras 2015), the only source of collocations for NNSs, has been available only since November 2017. So collocations can only be verified in corpora for Croatian as L1, which contain information about their frequency. For that reason collected collocations were verified in the *Veliki rječnik hrvatskoga jezika* (Anić 2003) and the hrWaC corpus (Ljubešić and Erjavec 2011), which is currently the largest Croatian language corpus, comprising 1.9 billion tokens. A total of 228 collocations were collected, of diverse frequency and associative strength and representing a range of grammatically and semantically possible collocations. Of these, 30 adjective-noun collocations were randomly chosen for the study. In this research, collocations in the hrWaC corpus were searched using CQL (Corpus Query Language) queries. Collocations were searched using the attribute 'word', so a CQL query for the collocation *osnovna škola* 'elementary school' looks like this:

[word="osnovn.*" & tag="A.*"] [word="škol.*" & tag="N.*"]

This query shows that the attribute value consists of "osnovna" and "škola" and that additional markers are used in order to make the query more precise. The equals sign (=) is used to add value to the attribute. The mark "tag" refers to a specific part of speech that is searched. Since this research focuses on adjective-noun collocations, A refers to adjectives and N refers to nouns. The wildcard (.*) represents different forms of search words, that is, encompasses adjectives' grammatical gender, number, and case, and nouns'

type, gender, number, and case. Finally, in order to point to two attribute values that are searched, the ampersand symbol (&) is used.[1]

Using computational linguistic tools in the hrWaC corpus, the frequency of collected adjective-noun collocations was calculated. It should be mentioned that CQL queries for each collocation did not contain case endings that would indicate the frequency of a specific case form; therefore wildcard (.*) was used to include all case forms of each collocation. Research has shown that some case forms appear more often than others in CFL; that is, a large percentage of nouns and adjectives does not appear equally in all cases (Cvikić and Jelaska 2003). For example, in traditional Croatian grammars, dative is listed in third and accusative in fourth place. However, relevant CFL research shows (Cvikić and Jelaska 2003; Cvikić and Jelaska 2007) that in terms of frequency accusative is one of the most important cases, whereas dative is significantly rare. So the order of teaching cases in CFL should be adapted accordingly. Because of that, frequency of collocations was determined regardless of the case form, rather than limiting it to specific case forms. The highest frequency in a row was considered as most relevant. For example, the most frequent form of the collocation *osnovna škola* was its genitive form *osnovne škole* with 32,168 occurrences. The frequency range of the 30 target collocations was set as follows: higher frequency (HF) from 1,222 to 43,106 occurrences and lower frequency (LF) from 15 to 953 occurrences in the hrWac corpus (Appendix 1).

Collocational strength was determined by using the associative connectedness assessment method, where participants assess the range of connection between the association and the collocational constituents (Maki 2007; Barrett and Fossum 2001). Finally, the associative strength of collocational constituents was assessed by 188 native speakers of Croatian, who estimated the strength of association between collocational constituents on a scale from 1 to 5. After this assessment, descriptive data (M, SD) that represent the average associative strength were calculated for each collocation. The average mean value (M) for the 30 target collocations was 3.56, so collocations were divided into two groups: all collocations with mean value (M) above 3.56 were collocations with strong associative strength (SAS), whereas the rest of the collocations with mean value (M) under 3.56 were labeled as collocations with weak associative strength (WAS). The associative strength range of the 30 target collocations was: strong associative strength with mean value (M) from 4 to 3.61 and weak associative strength with mean value (M) from 2.61 to 3.5 (Appendix 1).

[1] More detailed instructions on the use and purpose of CQL queries can be seen on the web page of Sketch Engine at https://www.sketchengine.co.uk/corpus-querying/ (Accessed Zagreb, 17 March 2017).

4.3. Tasks

The test consisted of three fill-in-the-blank tasks. In every task, a collocational base was given (a noun) and participants were asked to demonstrate their productive knowledge of the collocator (adjective). Before testing, two native speakers of Croatian who are also experts in CFL teaching were asked to review all three fill-in-the-blank tasks.

The first two of the three tasks were designed to examine the influence of morphological features. There were 40 isolated sentences without a wider context. In the first task, collocations were in the unmarked and independent nominative case, which means that the noun, as well as the adjective, was in the nominative case (example 1).

(1) Najnovije je istraživanje potvrdilo da su skup automobil, dizajnerski sat i vikendica najvažniji **statusni** simboli koji se koriste za javno pokazivanje moći i ugleda.

Recent research has confirmed that an expensive car, a designer watch, and a summer house are the most important **status** symbols used to publicly demonstrate power and reputation.

In the second task the same collocations from the task with collocations in the nominative case were given in oblique (or dependent) cases, which means that if the noun was in the instrumental case, so was the adjective (example 2). In other words, in the Croatian language adjectives agree with the noun they modify in case (nominative is the independent case and genitive, dative, accusative, locative, and instrumental case are oblique cases), gender (masculine, feminine, neuter), and grammatical number (singular or plural). Since this is one of the first studies of collocations in CFL, no exceptions to grammatical agreement between nouns and adjectives were included.

(2) Najnovije je istraživanje potvrdilo da su skup automobil, dizajnerski sat i vikendica među najvažnijim **statusnim** simbolima koji se koriste za javno pokazivanje moći i ugleda.

Recent research has confirmed that an expensive car, a designer watch, and a summer house are considered as the most important **status** symbols used to publicly demonstrate power and reputation.

The distribution of collocations from the tasks with collocations in the nominative case and oblique cases is given in Table 1 on the following page.

In the third part of the test, in order to examine the influence of wider context, frequency, and associative strength, participants were given a fill-in-the-blank task in which collocations were put into a more informative, wider

Table 1. Distribution of adjective-noun collocations according to frequency and associative strength in the tasks with collocations in the nominative case and oblique cases

Fill-in-the-blank tasks (nominative and oblique cases)	
HIGHER FREQUENCY + STRONG ASSOCIATIVE STRENGTH	LOWER FREQUENCY + STRONG ASSOCIATIVE STRENGTH
1. ljudska prava 'human rights'	1. sinkronizirano plivanje 'synchronized swimming'
2. društvena mreža 'social network'	2. samohrani roditelj 'single parent'
3. kulturna baština 'cultural heritage'	3. organska hrana 'organic food'
4. kreditna kartica 'credit card'	4. masovna proizvodnja 'mass production'
5. farmaceutska industrija 'pharmaceucital industry'	5. akcijski film 'action movie'
HIGHER FREQUENCY + WEAK ASSOCIATIVE STRENGTH	LOWER FREQUENCY + WEAK ASSOCIATIVE STRENGTH
1. službeni jezik 'official language'	1. statusni simbol 'status symbol'
2. zdravstvena zaštita 'medical insurance'	2. umjetničko klizanje 'figure skating'
3. sportske novine 'sport newspapers'	3. ručna torba 'hand bag'
4. nevladina organizacija 'non-governmental organization'	4. virtualni prijatelj 'virtual friend'
5. osobni podatak 'personal data'	5. stereotipna uloga 'stereotypical role'

context. For that reason, a topic about food and dietary habits among Croatians was chosen to compile a short text with adjective-noun collocations. All 20 collocations (base + collocator) in the wider context were in oblique cases. The task with collocations in the wider context includes 10 of the same collocators as do the tasks with collocations in the nominative case and oblique

cases (*ljudska prava, društvena mreža, farmaceutska industrija, samohrani roditelj, organska hrana, masovna proizvodnja, zdravstvena zaštita, nevladina organizacija, statusni simbol, virtualni prijatelj*). The other 10 collocations were new and included only in the task with collocations in the wider context (*prehrambene navike, plava riba, životni standard, smrtonosna bolest, ruralno područje, bijelo brašno, klimatski uvjeti, turistički centar, kupovna moć i kućni budžet*) (see Table 2).

Table 2. Distribution of adjective-noun collocations according to frequency and associative strength in the task with collocations in the wider context

Fill-in-the-blank tasks with wider context	
HIGHER FREQUENCY + STRONG ASSOCIATIVE STRENGTH	LOWER FREQUENCY + STRONG ASSOCIATIVE STRENGTH
1. ljudska prava 'human rights'	1. plava riba 'pelagic fish'
2. društvena mreža 'social network'	2. samohrani roditelj 'single parent'
3. prehrambene navike 'dietary habits'	3. organska hrana 'organic food'
4. životni standard 'standard of living'	4. masovna proizvodnja 'mass production'
5. farmaceutska industrija 'pharmaceutical industry'	5. smrtonosna bolest 'deadly disease'
HIGHER FREQUENCY + WEAK ASSOCIATIVE STRENGTH	LOWER FREQUENCY + WEAK ASSOCIATIVE STRENGTH
1. ruralno područje 'rural area'	1. statusni simbol 'status symbol'
2. zdravstvena zaštita 'medical insurance'	2. bijelo brašno 'white flour'
3. klimatski uvjeti 'climate conditions'	3. turistički centar 'tourist center'
4. nevladina organizacija 'non-governmental organization'	4. virtualni prijatelj 'virtual friend'
5. kupovna moć 'purchasing power'	5. kućni budžet 'household budget'

4.4. Procedure

The study was conducted in February 2016 in Krakow, Poland. In order to investigate the influence of morphological features, the tasks with collocations in the nominative case and oblique cases had the same adjective-noun collocations, and for that reason the participants were divided into two groups. The first group was asked to complete the tasks with collocations in the nominative and the task with collocations in a wider context, and the second group was asked to complete the tasks with collocations in the oblique cases and also the task with collocations in a wider context. Each task was presented separately and the order of tasks, as well as of collocations in the tasks was varied. Participation in this research was voluntary. The participants used a password and were not asked to reveal their personal identity or to provide contact information to the researchers. The coding of the collected data was done by the authors. Only adjective-noun collocations from Table 1 and Table 2 were accepted as correct answers, while all other possible answers were grouped into different types of answers. Namely, incorrect answers were grouped into three categories (lexical, grammatical, and overlapping answers), which were then divided into different types or deviations (e.g., wrong case of the adjective, wrong part of speech, approximation …), but the qualitative analyses of these answers is not the focus of this study. Since the emphasis of the study is on factors (morphological features, frequency, and associative strength) that influence collocational usage among Polish speakers of CFL, only the quantitative results are presented below.

5. Results

5.1. Answer Accuracy Analysis of Tasks with Collocations in the Nominative Case and Oblique Cases for Lower and Higher CFL Proficiency Level

In order to examine the influence of morphological features on productive knowledge of noun collocations in the nominative and oblique cases, descriptive data on the answer accuracy proportions were calculated for lower and higher proficiency participants (Table 3 on the following page). Due to rather small samples, the table shows medians (Mdn) and semi-interquartile ranges (Q).

Two nonparametric Mann-Whitney U tests, which are used for examining differences between independent groups, were performed in order to examine the differences in answer accuracy in productive tasks with isolated sentences with morphologically unmarked collocations in the nominative case and morphologically marked collocations in oblique cases for higher

Table 3. Descriptive data on answer accuracy proportions in productive tasks with isolated sentences with morphologically unmarked collocations in the nominative case and collocations in oblique cases among participants of lower and higher CFL proficiency level

			Median	Semi-interquartile ranges
Lower CFL proficiency	Nominative	(N = 8)	.25	.08
	Oblique cases	(N = 6)	.18	.08
Higher CFL proficiency	Nominative	(N = 6)	.45	.10
	Oblique cases	(N = 7)	.50	.005

and lower proficiency participants. The Mann-Whitney U test confirmed that there is no difference in answer accuracy between participants who solved the productive tasks with isolated sentences with morphologically unmarked collocations in the nominative case and those who solved the tasks with morphologically marked collocations in oblique cases, neither among lower proficiency participants ($z = .71$, $p = .477$) nor among higher proficiency participants ($z = -.50$, $p = .617$).

Since it was confirmed that answer accuracy among participants considering their proficiency level (lower, higher) is equal for tasks with collocations in the nominative case and oblique cases, in the following step of the analysis the data were grouped exclusively based on participants' CFL proficiency (lower level N = 14, higher level N = 13). The answer-accuracy proportions were calculated for each participant for four types of collocations in tasks with collocations in the nominative and oblique cases, considering their frequency and associative strength, as follows:

(i) collocations of higher frequency and strong associative strength
(ii) collocations of higher frequency and weak associative strength
(iii) collocations of lower frequency and strong associative strength
(iv) collocations of lower frequency and weak associative strength.

In order to examine the influence of frequency and associative strength on productive knowledge of noun collocations, a 2 × 2 × 2 ANOVA with collocational frequency (lower, higher) and associative strength (strong, weak) as within-participant factors and with CFL proficiency level (lower, higher) as

a between-participants factor was performed. The obtained results are presented in Table 4 below.

Table 4. Results of the 2 × 2 × 2 ANOVA with collocational frequency (lower, higher) and associative strength (strong, weak) as within-participant factors and with CFL proficiency level (lower, higher) as a between-participants factor for answer accuracy

Source of variation	df	F	MSE	*p*
CFL proficiency level	1.25	39.414	.056	.001
Frequency	1.25	106.703	.031	.0001
Associative strength	1.25	3.345	.026	.079
CFL proficiency level * frequency	1.25	9.786	.031	.004
CFL proficiency level * Associative strength	1.25	.820	.026	.374
Frequency * Associative strength	1.25	1.836	.018	.188
CFL proficiency level * Frequency * Associative strength	1.25	3.630	.018	.068

The main effects of CFL proficiency levels and frequency on answer accuracy were found. Furthermore, the two-way interaction of CFL proficiency level and collocational frequency on answer accuracy was found to be significant. Duncan's post hoc test showed that among higher CFL proficiency participants the answer accuracy for collocations of higher frequency is higher (M = .72; SE = .07) than for collocations of lower frequency (M = .26; SE = .03; $p < .001$). Among lower CFL proficiency participants, answer accuracy is on average significantly lower ($p < .01$) than among higher CFL proficiency participants. However, a pattern in the results was noticed: answer accuracy was higher for collocations of higher frequency (M = .32; SE = .07) than for collocations of lower frequency (M = .07; SE = .03; $p < .001$). These results are presented in Figure 1 on the following page.

5.2. Answer Accuracy Analysis of Tasks with Collocations in a Wider Context for Lower and Higher CFL Proficiency Level

In the next stage of the analysis, answer accuracy in the productive task with collocations in a wider context among participants of lower (N = 14) and higher (N = 13) CFL proficiency level was examined.

In order to investigate answer accuracy in the productive task with collocations in a wider context, a 2 × 2 × 2 ANOVA with collocational frequency

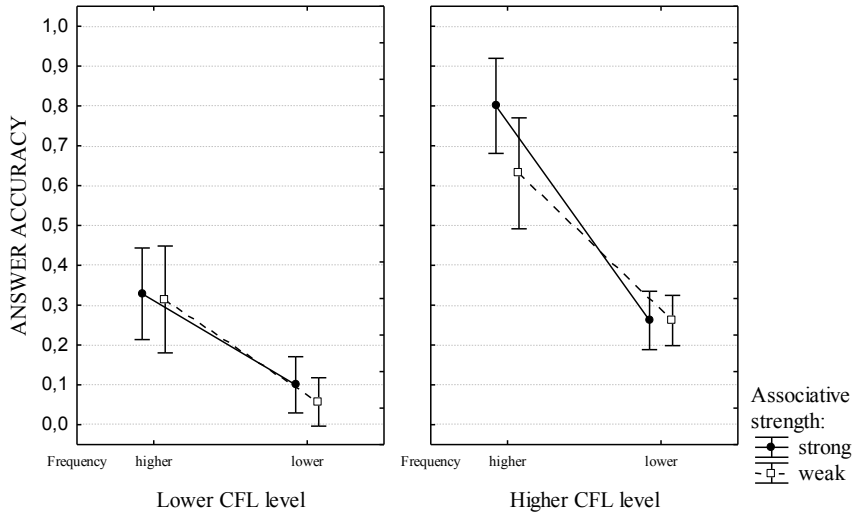

Figure 1. Answer accuracy in productive tasks with isolated sentences considering frequency (higher, lower) and associative strength (weak, strong) of collocational constituent among participants of lower and higher CFL proficiency.

(lower, higher) and associative strength (strong, weak) as within-participant factors and with CFL proficiency level (lower, higher) as a between-participants factor was performed. The results are shown in Table 5.

Table 5. Results of the 2 × 2 × 2 ANOVA with collocational frequency (lower, higher) and associative strength (strong, weak) as within-participant factors and with CFL proficiency level (lower, higher) as a between-participants factor for answer accuracy.

Source of variation	**df**	**F**	**MSE**	***p***
CFL proficiency level	1.25	23.404	.055	.001
Frequency	1.25	39.419	.035	.001
Associative strength	1.25	20.478	.017	.001
CFL proficiency level * Frequency	1.25	1.328	.035	.260
CFL proficiency level * Associative strength	1.25	.001	.017	.983
Frequency * Associative strength	1.25	106.607	.024	.0001
CFL proficiency level * Frequency * Associative strength	1.25	3.035	.024	.094

The main effect of CFL frequency level on answer accuracy was found. Answer accuracy was on average higher among higher CFL proficiency participants (M = .40; SE = .03) than among lower CFL proficiency participants (M = .18; SE = .03; $p < .001$). Furthermore, the effects of frequency and associative strength as well as the two-way interaction of frequency and associative strength on answer accuracy were found to be significant. Duncan's post hoc test confirmed that answer accuracy was the highest for collocations of higher frequency and strong associative strength (M = .61; SE = .05) compared to all other types of collocations (collocations of higher frequency and weak associative strength (M = .19; SE = .03; $p < .001$), collocations of lower frequency and strong associative strength, (M = .07; SE = .02; $p < .001$), collocations of lower frequency and weak associative strength (M = .27; SE = .03; $p < .001$). It can be noted that answer accuracy was unexpectedly higher for collocations of lower frequency and weak associative strength, compared to collocations of higher frequency and weak associative strength ($p < .05$), and lower frequency and strong associative strength ($p < .001$). These results are presented in Figure 2.

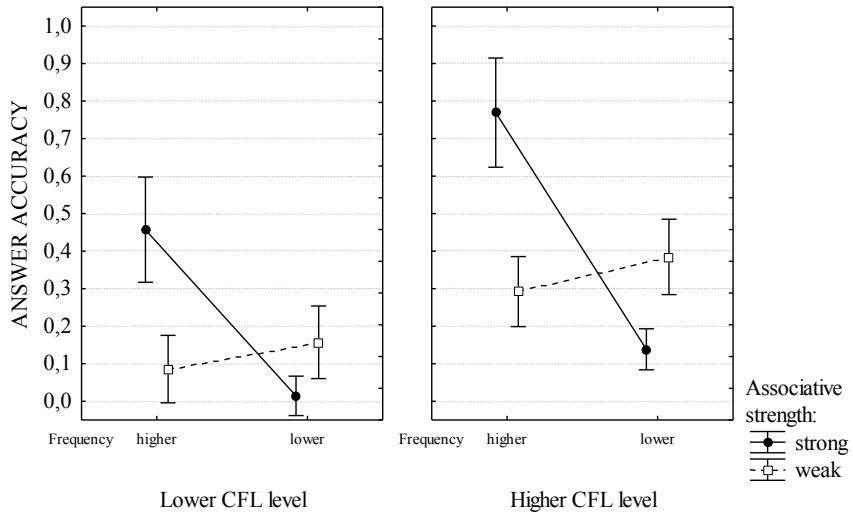

Figure 2. Answer accuracy in productive tasks with collocations in a wider context considering their frequency (higher, lower) and associative strength of collocational constituents (weak, strong) among respondents of lower and higher CFL proficiency level.

5.3. Discussion

In this study, productive knowledge of Croatian adjective-noun collocations among Polish students of Croatian was investigated, considering word frequency, associative strength, morphological features, and CFL proficiency level.

In order to investigate the influence of morphological features on productive knowledge of noun collocations among participants of lower and higher CFL proficiency, calculations of the differences in answer accuracy proportions were made. Hypothesis H1a, predicting that participants of lower CFL proficiency level would have better results in tasks with collocations in the nominative case, was not confirmed, while hypothesis H1b, for higher CFL proficiency, was confirmed. In other words, this analysis showed that participants at both proficiency levels know collocations in the nominative and oblique cases equally well. Even though Polish students of CFL can have difficulties with noun-adjective agreement (Aleksovski 2014; Kordić and Vidović Bolt 2013), it can be concluded that morphological features of lexical units do not affect collocational usage and that Polish CFL students had adequately acquired basic Croatian noun and adjectival declensions for adjective-noun collocations used in this research. The results mentioned for Polish students with higher CFL proficiency level further confirm prior research on collocations in heterogeneous groups, which suggested that the knowledge of collocations in the nominative and oblique cases is equal only at B2 proficiency level (Ordulj 2017). This is also in line with a study on collocational competence of native speakers of Croatian done by Ordulj and Cvikić (2017). On the other hand, morphological features seem to be a crucial factor in collocational usage among heterogeneous CFL learner groups of lower proficiency (Ordulj 2017), which stands in contrast to findings about Polish CFL students of lower proficiency level.

Apart from morphological features, in this research on productive collocational knowledge it was important to examine the influence of the frequency and associative strength of collocations on the usage of noun collocations among Polish CFL students of lower and higher proficiency levels. The analysis showed that hypothesis H2 is partially confirmed. Namely, higher-proficiency participants on average demonstrated higher answer accuracy compared to lower proficiency participants, but it can be observed that only frequency affected collocational usage in both groups. In fact, a pattern can be observed: There is higher answer accuracy for collocations of higher frequency, compared to collocations of lower frequency. These findings confirm that frequency of lexical co-occurrence in the linguistic input plays an important role in collocational acquisition among Polish CFL students for productive knowledge on both higher and lower proficiency levels. Since high-frequency collocations create stronger mental representations in the mental lexicon of

non-native speakers than low-frequency collocations do, these findings are in line with prior research (Siyanova-Chanturia and Spina 2015), including on the role of frequency in language acquisition (Ellis 2002, 2012; Jaén 2007; Nation 2001). Associative strength (i.e., the predictability of collocational constituents) did not appear to be significant in this study, even though it was found to be an important factor in collocational usage of B2 CFL learners with different first languages (Ordulj 2017).

Finally, one goal of this research was to investigate the influence of frequency and associative strength on productive knowledge of noun collocations in a wider context among lower and higher proficiency respondents. Hypothesis H3 is confirmed. In other words, higher proficiency participants demonstrate higher accuracy than lower proficiency participants. When it comes to answer types, regardless of the proficiency level, answer accuracy was highest for collocations of high frequency and strong associative strength, which are more frequent in non-native usage and are furthermore characterized by stronger predictability of collocational constituents which can be produced as wholes or sequences associated in the mental lexicon of the speaker (Nesselhauf 2003; Pawley and Syder 1983; James 1950). However, the analysis unexpectedly showed that Polish students of CFL have better knowledge of collocations with lower frequency and weak associative strength than of collocations of higher frequency and weak associative strength, which was also confirmed for heterogeneous learner groups (Ordulj 2017). As Blagus Bartolec (2014) points out, collocations are lexical units with a highly emphasized communicative function, and development of collocational competence is influenced by various factors. It is reasonable to assume, following González Fernández and Schmitt (2015: 114), that "L2 learners might better know those collocations which are likely to be encountered in daily situations, and therefore to have a more useful communicative function, compared to collocations whose function is more restricted to specific contexts". So a possible cause of the domination of collocations with lower frequency and weak associative strength among Polish learners of CFL could be exposure, that is, repeated usage of these collocations in everyday life through the personal interests of the participants (social networks, traveling, food, etc.) or the influence of L1. Since corpus frequency relates to acquisition of collocations only moderately, as was indicated by González Fernández and Schmitt (2015), it would be worth investigating and comparing corpus frequency as indicated by hrWaC with intuitive frequency assessment by NSs and NNSs of Croatian, which could deepen the insights of this research. However, in order to get more precise answers, further research on the influence of context and frequency on collocational usage in CFL is needed.

Since the main goal of the present study was to investigate factors influencing the productive knowledge of adjective-noun collocations among Polish

CFL students, a qualitative analysis of collocational errors will not be presented in this paper.

6. Conclusion

Based on the analysis conducted, the following conclusions can be drawn about factors influencing productive collocational knowledge among Polish students of CFL.

Firstly, participants at both proficiency levels demonstrate that morphological features of lexical units do not influence production of collocations in CFL.

Secondly, frequency is a factor that influences production of collocations at both CFL proficiency levels, whereas associative strength does not appear to be significant.

Thirdly, production of collocations in a wider context among participants at both CFL proficiency levels is influenced by both frequency and associative strength, whereas the influence of context falls for further research.

In general, it can be concluded that the acquisition of adjective-noun collocations among Polish students of CFL is satisfactory. However, answers to research problems and hypotheses examined in this study were not unambiguous, marking it as just a beginning of research on collocational competence among Polish students of CFL. It should be mentioned that there are some limitations to this research method that should be taken into account. First of all, it was conducted on a rather small number of participants; therefore the results should not be generalized. Furthermore, in order to get better insight into the overall acquisition of collocations in CFL, additional research with different collocational types is required, which would encompass both heterogeneous and homogeneous CFL learner groups.

References

Aleksovski, Marinela. (2014) "Poljski" hrvatski". Marica Čilaš Mikulić and Antonio-Toni Juričić, eds. *CROATICUM—Savjetovanje za lektore hrvatskoga kao inoga jezika*. 2 zbornik radova. Zagreb: FF-press, 11–28.

Anić, Vladimir. (2003) "Veliki rječnik hrvatskoga jezika". Zagreb: Školska knjiga.

Barrett, Lisa Feldman and Thyra Fossum. (2001) "Mental representations of affect knowledge". *Cognition and emotion* 15: 333–63.

Begagić, Mirna. (2014) "English language students' productive and receptive knowledge of collocations". *Explorations in English language and linguistic*, 2(1): 46–67. Available at: https://hrcak.srce.hr/139436.

Benson, Morton. (1985) "Collocations and Idioms". Robert Ilson, ed. *Dictionaries, lexicography, and language learning*. Oxford: Pergamon Press, 61-68.

Blagus Bartolec, Goranka. (2014) *Riječi i njihovi susjedi. Kolokacijske sveze u hrvatskom jeziku*. Zagreb: Institut za hrvatski jezik i jezikoslovlje.

Bonci, Angelica. (2002) "Collocation restrictions in Italian as a second language: A case control study". *Tuttitalia* 26: 3–14.

Borić, Neda. (1998) "Semantički aspekt kolokacijskih odnosa s kontrastivnog stajališta". *Strani jezici* 27(2): 72–79.

————. (2002) "Leksikografska prezentacija kolokacijskih sklopova". *Strani jezici* 31(1–2): 55–63.

————. (2004) "Kolokacije kao dio leksičkog pristupa u nastavi stranih jezika". Nada Ivanetić, Boris Pritchard, and Diana Stolac, eds. *Suvremena kretanja u nastavi jezika. Zbornik HDPL*. Zagreb-Rijeka: Graftrade, 63–68.

Burić, Helena and Josip Lasić. (2012) "Kolokacije u nastavi inojezičnoga hrvatskog". Renata Relja, ed. *Zbornik radova Filozofskoga fakulteta u Splitu 5*. Split: Filozofski fakultet, 233–49.

Cvikić, Lidija and Marija Bošnjak. (2004) "Pogled u obilježja i probleme učenja riječi u hrvatskome kao nematerinskome jeziku". Nada Ivanetić, Boris Pritchard, and Diana Stolac, eds. *Suvremena kretanja u nastavi jezika. Zbornik HDPL*. Zagreb-Rijeka: Graftrade, 111–21.

Cvikić, Lidija and Zrinka Jelaska. (2003) "Poučavanje imeničke sklonidbe u nastavi hrvatskoga kao stranoga jezika". Nada Ivanetić, Boris Pritchard, and Diana Stolac, eds. *Psiholingvistika i kognitivna znanost u hrvatskoj primijenjenoj lingvistici. Zbornik HDPL*. Zagreb-Rijeka: Graftrade, 167–77.

————. (2007) "Morfološka raslojenost imenica u svjetlu inojezičnoga hrvatskoga". *Riječ* 13(2): 66–78.

Duplančić Rogošić, Gorana. (2007) *Obrada kolokacija u englesko-hrvatskim, hrvatsko-engleskim i englesko-engleskim rječnicima*. MA thesis, Zadar Filozofski fakultet.

El-Dakhs, Diana. (2015) "The lexical collocational competence of Arab undergraduete EFL learners". *International journal of English linguistics* 5(5): 60–74. Available online: https://www.researchgate.net/profile/Dina_ElDakhs/publication/282364488_The_Lexical_Collocational_Competence_of_Arab_Undergraduate_EFL_Learners/links/560e9d9e08aec422d111d7db/The-Lexical-Collocational-Competence-of-ArabUndergraduate-EFL-Learners.pdf.

Ellis, Nick C. (2002) "Frequency effects in language processing". S*tudies in second language acquisition* 24: 143–88.

————. (2012) "What can we count in language, and what counts in language acquisition, cognition, and use?" Stephan Th. Gries and Dagmar S. Divjak, eds. *Frequency effects in language learning and processing*. Vol. 1. Berlin: Mouton de Gruyter, 7–34.

Gitsaki, Christina. (1999) *Second language lexical acquisition. A study of the development of collocational knowledge*. San Francisco-London-Bethesda: International Scholars Publications.

González Fernández, Beatriz and Norbert Schmitt. (2015) "How much collocation knowledge do L2 learners have?". *International journal of applied linguistics* 166(1): 94–126. Available at: http://doi.org/10.1075/itl.166.1.03fer.

Hill, Jimmie. (1999) "Collocational competence". *English teaching professional* 11: 4–7.

————. (2000) "Revising priorities: From grammatical failure to collocational success". Michael Lewis, ed. *Teaching collocation: Further developments in the lexical approach*. London: Language Teaching Publications, 49–50.

Ivir, Vladimir. (1992–1993) "Kolokacije i leksičko značenje". *Filologija* 20/21: 181–89.

Jaén, María Moreno. (2007) "A corpus-driven design of a test for assessing the ESL collocational competence of university students". *International journal of English studies* 7(2): 127–47. Available at: http://revistas.um.es/ijes/article/view/49031/46901.

James, William. (1950) *The principles of psychology*. Vol. 1. New York: Dover Publications.

Kordić, Ana and Ivana Vidović Bolt. (2013) "Važnost kontrastivnog pristupa u poučavanju stranoga jezika-primjer ovladavanja hrvatskim za polj ske govornike". Available at: https://repozytorium.amu.edu.pl/jspui/bitstream/10593/9941/1/Kodric%20Ana,%20Vidovic%20Bolt%20Ivana.pdf.

Koya, Taeko. (2003) "A study of collocation in English and Japanese noun-verb combinations". *Intercultural communication studies* 12(1): 125–45.

Lewis, Michael. (1997) *Implementing the lexical approach: Putting theory into practice*. Hove, London: Language Teaching Publications.

Ljubešić, Nikola and Tomaž Erjavec. (2011) "hrWaC and slWac: Compiling web corpora for Croatian and Slovene". Ivan Habernal and Vàclav Matoušek, eds. *Text, speech, and dialogue. TSD 2011. Lecture notes in computer science*. Vol. 6836. Berlin: Springer, 395–402.

Maki, William. (2007) "Judgments of associative memory". *Cognitive psychology* 54(4): 319–53. Available at: https://pdfs.semanticscholar.org/773d/8c08d0a3733e39002522dc37404f7fc87508.pdf.

Martyńska, Małgorzata. (2004) "Do English language learners know collocations? " *Investigationes linguisticae* 11: 1–12. Available at: http://pressto.amu.edu.pl/index.php/il/article/view/9049/8808.

Mikelić Preradović, Nives, Monika Berać, and Damir Boras. (2015) "Learner corpus of Croatian as a second and foreign language". Kristina Cergol Kovačević and Sanda Lucija Udier, eds. *Multidisciplinary approaches to multilingualism*. Frankfurt am Main: Peter Lang: 107–26.

Miščin, Evelina. (2012) *Glagolske kolokacije u engleskome jeziku medicinske struke*. PhD dissertation, Osijek Filozofski fakultet. Available at: http://dar.nsk.hr/?vdoc=8794&page=1.

Nation, Paul. (2001) *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Nesselhauf, Nadja. (2003) "The use of collocations by advanced learners of English and some implications for teaching". *Applied linguistics* 24(2): 223–42.

Ordulj, Antonia. (2016) *Kolokacije kao sastavnica leksičke kompetencije u hrvatskoma kao inom jeziku*. PhD dissertation, Zagreb Filozofski fakultet. Available at: http://darhiv.ffzg.unizg.hr/7358/1/Ordulj_Kolokacije%20kao%20sastavnica%20leksicke%20kompetencije%20u%20HIJ-u_doktorski%20rad_2016.pdf.

————. (2017) "Kolokacijska kompetencija u hrvatskom kao inom jeziku". *Philological studies* XV(2): 196–218. Available at: http://philologicalstudies.org/dokumenti/2017/issue2/17_Ordulj.pdf.

Ordulj, Antonia and Lidija Cvikić. (2017) "Čimbenici ovladanosti kolokacijama kod izvornih govornika hrvatskoga jezika". *Croatian journal of education* 19(4): 1167–97. Available online: https://doi.org/10.15516/cje.v19i4.2655.

Pawley, Andrew and Frances H. Syder. (1983) "Two puzzles of linguistic theory: Nativelike selection and nativelike fluency". Jack C. Richards and Richard W. Schmidt, eds. *Language and communication*. London: Longman, 191–225.

Petrović, Bernardina. (2007) "Razvijanje kolokacijske kompetencije u hrvatskome kao stranom i drugom jeziku". *Strani jezici* 3: 31–38.

————. (2008a) "Nacrtak za Kolokacijski rječnik hrvatskoga jezika". Marko Samardžija, ed. *Vidjeti Ohrid: Referati hrvatskih sudionica i sudionika za XIV. Međunarodni slavistički kongres*. Zagreb: Hrvatsko filološko društvo—Hrvatska sveučilišna naklada, 225–35.

————. (2008b) "Glagoli emocionalnih stanja u kolokacijskim strukturama i leksikografskom opisu". Ines Srdoč Konestra and Silvana Vranić, eds. *Riječki filološki dani 7: Zbornik radova s Međunarodnoga znanstvenog skupa Riječki filološki dani održanog u Rijeci od 16. do 18. studenoga 2006*. Rijeka: Filozofski fakultet, 589–99.

Podboj, Martina. (2013) "Red riječi u hrvatskom i poljskom jeziku: Poteškoće s kojima se susreću poljski student kroatistike pri usvajanju reda riječi u hrvatskom jeziku". *RASPRAVE Časopis Instituta za hrvatski jezik i jezikoslovlje* 39(2): 581–92.

Prichard, Boris. (1998) "O kolokacijskom potencijalu rječničkog korpusa". *Filologija* 30/31: 285–99.

Schmitt, Norbert. (2004) *Formulaic sequences: Acquisition, processing, and use*. Amsterdam: John Benjamins.

Shin, Dongkwang. (2007) "What collocations would be unpredictable for Korean EFL learners?" *Korean journal of applied linguistics* 32(2): 83–98.

Siyanova-Chanturia, Anna and Stefania Spina. (2015) "Investigation of native speaker and second language learner intuition of collocation frequency". *Language learning* 65(3): 533–62.

Sokolić, Nikolina and Ivana Vidović Bolt. (2012) "Pada li Poljaku lišće u listopadu? O hrvatsko-poljskim lažnim prijateljima". Marica Čilaš Mikulić, Antonio-Toni Juričić, and Sanda Lucija Udier, eds. *Croaticum—Savjetovanje za lektore hrvatskoga kao inoga jezika*. I zbornik radova. Zagreb: FF Press, 15–26.

Stojić, Aneta. (2012) *Kolokacije. Prilog teoriji i praksi*. Rijeka: Filozofski fakultet.

Štefić, Lidija, Marinka Mravak-Stipetić, and Vesna Borić. (2010) "Kolokacije u jeziku stomatologije: Primjeri iz oralne medicine". *Acta stomatologica Croatica* 44(3): 176–87. Available online: https://hrcak.srce.hr/58754.

Turk, Marija. (2000) "Višečlani izrazi s desemantiziranom sastavnicom kao nominacijske jedinice". Diana Stolac, ed. *Riječki filološki dani. Zbornik radova 3*. Rijeka: Filozofski fakultet, 477–86.

————. (2010) "Status kolokacija u hrvatskim jednojezičnim i dvojezičnim rječnicima". Lada Badurina and Danijela Bačić-Karković, eds. *Riječki filološki dani: Zbornik radova sa 6. međunarodnoga znanstvenog skupa Riječki filološki dani*. Rijeka: Filozofski fakultet, 537–47.

Faculty of Humanities and Social Sciences
Ivana Lučića 3
10 000 Zagreb
antoniasvetic@gmail.com
sokolicn@gmail.com

Appendix 1. List of Target Collocations With Frequency and Associative Strength Information

Collocation	M	SD	Frequency
ljudsko pravo	4	1.05	43.106
zdravstvena zaštita	3.5	1.14	16.895
osobni podatak	3.48	1.15	16.619
kulturna baština	4.39	0.81	14.524
društvena mreža	4.24	0.95	11.696
nevladina organizacija	3.14	1.27	4.521
kreditna kartica	4.57	0.72	4.258
prehrambena navika	3.67	1.12	3.048
životni standard	3.93	1.04	2.769
ruralno područje	3.46	1.11	2.078
službeni jezik	3.32	1.16	1.887
sportske novine	3.47	1.08	1.455
farmaceutska industrija	3.98	1.02	1.446
kupovna moć	3.07	1.23	1.341
klimatski uvjet	3.13	1.17	1.222
statusni simbol	2.69	1.27	953
akcijski film	3.82	1.03	940
kućni budžet	3.39	1.2	836
samohrani roditelj	4.15	0.96	806
masovna proizvodnja	3.9	0.98	664
smrtonosna bolest	4.06	0.99	501
bijelo brašno	3.44	1.31	486
organska hrana	3.61	1.17	446
ručna torba	3.04	1.22	426
plava riba	3.68	1.37	423
umjetničko klizanje	3.47	1.24	364
turistički centar	3.26	1.15	333
sinkronizirano plivanje	3.76	1.23	249
virtualni prijatelj	2.86	1.24	183
stereotipna uloga	2.61	1.14	15

The Effects of Animacy and Givenness on Object Order in Croatian Child Language

Marta Velnić

Abstract: This study investigates how givenness and animacy influence object order (IO-DO vs. DO-IO) in ditransitive constructions in Croatian child language. We have conducted an elicitation task with 59 monolingual Croatian children (mean age = 4;4) and 36 adult controls (mean age = 21), in which the participants were asked to describe images depicting ditransitive actions. These actions differed with regard to givenness (DO given, or IO given) and animacy (IO animate, or both IO and DO animate). Both groups demonstrated an animacy effect, manifested as a significant increase of DO-IO productions when both objects were animate, compared to when only the IO was; adults presented DO-IO preference at ceiling level. Givenness had a statistically significant effect (p-value < 0.01) only in adults, but both groups were affected by the givenness of the DO. This paper supplies arguments to support previous indications that (1) DO-IO is the underlying order in Croatian ditransitives, and (2) that children do not have an IO-DO preference as has been reported by previous studies conducted on case-marking languages.

1. Introduction

This study investigates how a semantic and a pragmatic factor, animacy and givenness respectively, are reflected in object order in ditransitive structures in Croatian pre-schoolers and adults. In ditransitive structures in Croatian, both object orders—indirect-direct object (IO-DO) and direct-indirect object (DO-IO)—are grammatical and attested. However, word order is sensitive to animacy and givenness, and thus IO-DO and DO-IO are used in different contexts. According to the principles of information structure, the animate argument should precede the inanimate argument, while given (old) information should come before new information (Birner and Ward 2009). IO-DO has been found to be frequently produced in corpus data by adult speakers (Velnić 2018, forthcoming), but adult speakers of Croatian were nevertheless found to have a preference for DO-IO when animacy and givenness were controlled for (Velnić 2019).

Journal of Slavic Linguistics 27(1): 85–125, 2019.

We have tested 59 monolingual Croatian children (mean age = 4;4) and 36 adult controls (mean age = 21) with an elicitation task in which ditransitive sentences were elicited through images. The givenness of the recipient (IO) and the theme (DO) were manipulated throughout the occurrence of the respective referents in the target pictures, and reinforced with pictures of the same referent, before presenting the next target picture. The subject and IO were always animate, while the animacy of the DO was manipulated (inanimate vs. animate). This setup provided two animacy conditions: the prototypical animacy condition (IO-animate and DO-inanimate), and the balanced animacy condition (both objects animate).

The results show that both children and adults were sensitive to animacy, as the occurrence of DO-IO order was considerably increased when both objects were animate, compared to the condition in which only the IO was animate. Givenness was not found to be a statistically significant factor in the child data, but it had an effect on the adults' productions. This givenness effect was confined to the condition of the given DO, as the givenness of the IO did not affect object order production. A closer look at the child data also suggests that children pay more attention to the givenness of the DO. We found a marginal significance when applying a test different than for the adults. Children displayed a new > given preference in the prototypical animacy condition, but the trend of responses changed when the animacy was balanced and showed an inclination towards given > new.

The paper is structured as follows: in the background section, we describe ditransitive structures in Croatian, and provide a summary of the animacy-first order and the given-before-new principle, along with previous child language studies. Next, we formulate our research questions and lay out the predictions. The methodology and the results sections follow. In the discussion section, the results are examined in relation to our predictions. A brief summary concludes the paper.

2. Background

In this section we outline the literature necessary for formulating the research questions and predictions. We focus on ditransitive structures, animacy, and givenness.

2.1. Ditransitive Structures

Ditransitive structures are comprised of three arguments: the subject, the direct object (DO), and the indirect object (IO). The main interest of this study is the relationship between the DO and the IO. Various languages have different strategies for arranging the two objects. If a language has overt case marking,

such as Croatian, both object orders are possible, as shown in example (1). Croatian uses the accusative case to express the theme (DO) and the dative case to express the recipient (IO). Languages with no overt case marking, like English, have two different structures used to convey the different object orders, like in example (2).

(1) a. Marlon daje Stigu jabuku.
Marlon$_{NOM}$ give$_{PRES.3SG}$ Stig$_{DAT}$ apple$_{ACC}$
'Marlon is giving Stig an apple.'

b. Marlon daje jabuku Stigu.
Marlon$_{NOM}$ give$_{PRES.3SG}$ apple$_{ACC}$ Stig$_{DAT}$
'Marlon is giving an apple to Stig.'

(2) a. Marlon gave Stig an apple. (Double Object Dative—DOD)

b. Marlon gave an apple to Stig. (Prepositional Dative—PD)[1]

Corpus data have revealed that both adults and children use IO-DO more frequently than DO-IO (Velnić forthcoming).

A structural variation in ditransitives is present in a very limited portion of the Croatian lexicon, appearing only with three verbs: (*po*)*nuditi* 'offer', (*po*)*služiti* 'serve' and *pokloniti* 'give as a gift' (Zovko-Dinković 2007). The alternative expresses the recipient with the accusative and the theme with the instrumental case. An example of the two structures using 'offer' is presented in example (3).

(3) a. Marlon je ponudio Stigu jabuku.
Marlon$_{NOM}$ AUX offer$_{PST.SG}$ Stig$_{DAT}$ apple$_{ACC}$

a′ Marlon je ponudio jabuku Stigu.
Marlon$_{NOM}$ AUX offer$_{PST.SG}$ apple$_{ACC}$ Stig$_{DAT}$

b. Marlon je ponudio Stiga jabukom.
Marlon$_{NOM}$ AUX offer$_{PST.SG}$ Stig$_{ACC}$ apple$_{INS}$

b′ Marlon je jabukom ponudio Stiga
Marlon$_{NOM}$ AUX apple$_{INS}$ offer$_{PST.SG}$ Stig$_{ACC}$
'Marlon offered an apple to Stig.'

[1] Croatian also allows PP constructions, but only in cases where it is intended as a change of location, and not a change of possession (e.g., *Ivan je bacio loptu prema Ani* 'John$_{NOM}$ threw ball$_{ACC}$ towards Anna$_{DAT}$', in which case we do not expect Anna to catch the ball). These structures are not elicited in the current task but have occurred when both objects were animate, most likely due to an interpretation of a caused motion (Levin 2008). These occurrences were excluded due to the weight of the PP.

This possibility of case alternation is why 'offer' was chosen as one of the verbs to be elicited in our task.[2]

Gračanin-Yuksek (2006) includes three word orders in her analysis by also taking into consideration the location of the verb (V). We have displayed the VID order in (1a) and VDI in (1b); the third order analyzed by Gračanin-Yuksek (2006) is when the IO precedes the verb (*Marlon Stigu$_{DAT}$ daje loptu$_{ACC}$*). Gračanin-Yuksek (2006) suggests that VDI (1b) and IVD are base-generated orders, while VID (1a) is structurally ambiguous. Thus, both IO-DO (IVD) and DO-IO (VDI) are underlying under this analysis. However, a contextual approach used in Velnić (2019) has found that DO-IO is strongly preferred when the two factors are controlled for.

Conversely, as we will see in section 2.3 regarding the literature review on ditransitive structures in child language, children have been shown to have an IO-DO preference (Höhle et al. 2014; Mykhaylyk, Rodina, and Anderssen 2013), which differs from what has been found for the adults of the respective languages (Røreng 2011; Titov 2017) and Croatian adults (Velnić 2019). The aim of our study is to control for animacy and givenness and observe the effect they have on object order, as well as to determine which is the most used order when these two factors are neutral.

2.2. The Effect of Animacy on Word Order and its Acquisition

As previously mentioned, the animacy of a referent does not vary based on the context of discourse: if a referent designates an animate being, it will be animate, regardless of whether it has already been given or whether it is in focus. It is a semantic, not a pragmatic, property that shapes information structure. Animate entities are conceptually highly accessible and thus easier to retrieve (Branigan, Pickering, and Tanaka 2008). Animate entities are also more likely to be prominent in the discourse because discourse prominence is related to the speakers' empathy, and animate entities are more eligible than inanimate entities to be prominent (Malchukov 2008).

There is a vast body of research that indicates that animacy influences word order in the direction of animacy-first, which means that animate arguments precede inanimate ones. However, few studies have investigated it in relation to ditransitives, which is our focus here.

Kempen and Harbusch (2004) conducted a corpus study on German di transitive sentences. In German the theme and the recipient are marked as in Croatian. The authors checked the order of each of the possible pairs of gram-

[2] The alternating structure (accusative-instrumental) in example (3b) failed to be elicited in the children, most likely due to the low frequency of this structure; the adult controls had only produced it twice in the task. Thus, this structure is disregarded for the rest of the analysis.

matical functions included in a ditransitive structure (S & DO, S & IO, DO & IO) in relation to animacy, and they found a direct influence of animacy on word order: An inanimate IO was unlikely to precede the subject, but when both subject and IO were animate, the distribution of S-IO and IO-S was at chance level. This observation was made for the subject and the IO in ditransitives, as the DO was not animate.

In ditransitive structures, animacy is closely linked to the IO, as prototypically the recipient is animate, and the theme is not (henceforth, prototypical animacy). Thus, the IO should be in a privileged position, appearing as the first object. However, if animacy were the only factor at play, we would rarely see realisations of the DO-IO order in any language. But that is not what happens, as DO-IO occurrences were found in Croatian corpora (Kovačević 2004; Kuvač Kraljević and Hržica 2016), albeit to a lesser extent than IO-DO.

Studies on animacy in child language suggest that animacy is acquired rather early, since children from around the age of two are able to distinguish animate from inanimate NPs in an adult-like manner (de Marneffe et al. 2012). Like in adults, an obvious effect of animacy is noticed in the studies of active/passive use, with preference for passive sentences when only the patient is animate (Lempert 1989).

With regard to the effect on ditransitive structures, Cook (1975) conducted an act-out task with a wide age range of English-speaking children (ages 5–10), and presented evidence that the comprehension of ditransitive sentences is better when the animacy is prototypical than when it is not.[3] Moreover, both configurations with unbalanced animacy (IO-animate/DO-inanimate and DO-animate/IO-inanimate) were better comprehended than the constructions with balanced animacy (both objects animate, or both inanimate).

Snyder (2003: 56) has shown that young children (around the age of three) are very attentive to animacy in their choice of ditransitive structure and rely less on animacy as they grow older. Snyder's (2003) corpus data (from English and Tahitian French) suggest that, as children rely less on animacy, other factors influence their word order choices. She argues that children use animacy as a stand-in for information status, until they are able to grasp what constitutes given information for the interlocutor. The fact that animacy is more relevant at a young age suggests that there will be a difference between children and adults regarding the relevance of this factor in determining word order.

[3] The configuration of inanimate IO and animate DO was constructed by a simple rotation of the 'giving' relation of the test objects, such as 'give the man to the book', a sentence that would have been very unlikely outside the experimental setting.

2.3. The Effect of Givenness on Word Order and its Acquisition

Many languages are subject to the given-before-new principle (henceforth given > new), which entails that if all other factors are equal speakers will prefer to place the information that is familiar to the listener first and place the new information later in the sentence (Birner and Ward 2009).

The given > new principle originated for the Slavic languages with the Prague school of linguistics (Firbas 1964), and the effects of this factor are still debated. More precisely, divergent implications were made on how strict the principle is in the case of Czech: strict (Kučerová 2012) or less strict (Šimík, Wierzba, and Kamali 2014). Kučerová (2007) suggests that in Czech only SVO, the basic word order, can be used in a variety of contexts, while other word orders must be used only in contexts that relate to the givenness values of their elements. In Kučerová (2012), the research is expanded to Russian and Serbo-Croatian;[4] her claim is that in these languages givenness is always marked, with given elements preceding new ones, and a new > given order is argued to be ungrammatical. The analysis provided by Šimík, Wierzba, and Kamali (2014) for Czech is less strict, and the authors claim that given objects can occur anywhere in the sentence, excluding the final position, which receives default main sentence stress.

More specifically for Croatian, Velnić (2019) found a givenness effect in an acceptability-judgment task on word order choice conducted on adult speakers. In this experiment, IO-DO structures were considered more acceptable when the IO was given, while the DO-IO order was judged better in conditions when the DO was given or when neither object was given. Note that both of these orders are perfectly well-formed and their acceptance depended solely on the context which they appeared in. Conversely, the data from Velnić (2014) indicate that IO-DO is predominant in oral communication, with many fewer cases of the DO-IO order being attested (child directed speech: 60/304 occurrences were DO-IO; children: 19/258 occurrences were DO-IO). Velnić (forthcoming) analysed a portion of these data and found limited occurrences of new > given in the child data (2/12 of DO-IO occurrences); the adult data displayed only the given > new order.

Ditransitive structures can accommodate given > new with the DO-IO order when the theme is given, and with the IO-DO order when the recipient is given. Clifton and Frazier (2004) and Brown, Savova, and Gibson (2012) (for English) along with Kizach and Balling (2013) (for Danish) have shown that having a given > new order facilitates sentence processing for DOD but not for the PD—examples (2a) and (2b) above. It has been suggested that discourse information is incorporated into the structure of the DOD, but not into that of the PD, and thus the DOD has constraints on how the given and new infor-

[4] Kučerová's (2012) term.

mation is ordered, allowing only for given > new (Brown, Savova, and Gibson 2012). Kizach and Mathiasen (2013) have also found that native Polish speakers learning Danish as a second language acquire the native Danish pattern quickly, implying that Polish has the same givenness asymmetries between DOD and PD as Danish. In languages that do not have different structures for dative alternation, such as German and Russian, it has been found that DO-IO is the canonical order (Røreng 2011 for German; Titov 2017 for Russian), due to its wide contextual applicability, while the IO-DO is either contextually motivated (Røreng 2011) or signals a meaning not available to the DO-IO (Titov 2017). While both of these studies dealt with the background/focus distinction rather than the given/new distinction, their findings are still applicable in terms of which word order is the underlying one.

Studies conducted on the effect of givenness on child language have reached divergent results, and there is still no general consensus regarding the age when givenness is in place. According to Schaeffer and Matthewson (2005), children have difficulty with implementing givenness in word order, because they lack a pragmatic concept that allows them to systematically distinguish between their own beliefs and the beliefs of the interlocutor. They refer to this as the concept of Non-Shared Assumptions.[5]

However, there is a clear division in the research conducted on the givenness effect in ditransitive sentences, and it is dependent on some key characteristics of the target language: whether the language has dative alternation, that is, two syntactic structures such as the English double object dative and prepositional dative (example 2a–b), or whether it has case marking, for example accusative for the theme and dative for the recipient like Croatian. Studies on languages with dative alternation have found an effect of givenness, whereas studies in the latter group have found a preference for IO-DO. We will provide a description of each of these studies in turn.

One of the studies on a dative alternating language, English, has already been mentioned in section 2.2 with regard to its results on animacy: in a corpus study Snyder (2003) found a progressive effect of givenness on word order in ditransitive sentences. Before the age of seven, the givenness effect is noticeable, but other factors—such as animacy and weight—are more important in determining word order, and the corpus even contains new IOs being placed before the DO at ages six and seven (Snyder 2003: 53). At age seven, givenness becomes the most relevant factor for object placement, but the children are not adult-like yet. The author does not state explicitly in which proportion the two object orders are attested in the corpus, so we cannot conclude which word order is preferred.

A clearer effect of givenness was obtained by Stephens (2015) with elicited production tasks. She found that four-year-olds tend to produce given > new

[5] Referred to also as Non-Shared Knowledge in Schaeffer (1999).

orderings in their dative constructions. In conditions with given themes (DO), children consistently produced the PD (DO-IO order); when the recipient (IO) was given, the participants were more likely to produce a DOD (IO-DO order) (2015: 416). The same pattern was found in the adult controls (2015: 424). This is consistent with the studies on adult language referred to above, which found a stronger givenness effect on given themes compared to given recipients (Clifton and Frazier 2004; Kizach and Balling 2013).

Anderssen et al. (2014) conducted a semi-spontaneous production task on Norwegian children (ages 4–6). Like English, Norwegian exhibits the DOD and PD distinction. The authors find a givenness effect: the theme-given context yielded the PD structure most of the time, while the recipient-given condition was divided among PD and DOD productions, with the latter still being produced much more than in the theme-given conditions.

Among the studies that found a preference for IO-DO, Mykhaylyk, Rodina, and Anderssen (2013) analysed the distribution of IO-DO/DO-IO in ditransitive structures in Russian and Ukrainian three- to six-year-olds. The responses with no omissions were mostly expressed in the IO-DO order with very little variation across the two givenness conditions. Nonetheless, there was an observable difference with age, as the older children used more DO-IO in the theme-given condition, but IO-DO was still the generally preferred object order. This suggests that Russian and Ukrainian children did not integrate the context in their ditransitive productions.

Höhle et al. (2014) conducted a test on German five-year-olds in which they checked how faithfully the children reproduced ditransitive structures that violated word order (*ACC-DAT)[6] or definiteness (*indef-def) constraints. They found that children faithfully reproduced sentences with no violations, but in the case of violations they reproduced definiteness violations more readily than word order violations. This means that they faithfully reproduced the constraint-respecting IO-DO sentences, but the constraint-violating DO-IO sentences were also often reproduced as IO-DO. This shows that keeping IO-DO is more relevant than having the definite NP precede the indefinite NP. Givenness is not identical to definiteness, but they are related properties, as the given argument can be expressed with a definite NP while a new argument is not likely to be expressed with a definite NP. However, the target sentences were provided in isolation, and a wider context might have strengthened the givenness effect as opposed to only marking it with a definite/indefinite article.

None of the studies above balanced animacy, using only the prototypical animacy configuration: IO-animate and DO-inanimate.

Croatian is like Russian, Ukrainian, and German regarding how the theme and recipient are marked. The predominance of the IO-DO order is

[6] They assume that IO-DO is the unmarked order.

evident from the corpus data, as an analysis of the Double Object DataBase (Velnić 2014), based on (Kovačević 2004), shows a predominant use of the IO-DO order in both children and child-directed speech. Velnić (forthcoming) analysed this database and found that children use both given > new and new > given word orders. However, the corpus data had limited instances of combinations of given and new objects, as most of the objects were accessible. An overview of these studies is provided in Table 1 on page 94.

Since Croatian marks the theme and the recipient like the languages in the latter group, i.e., by case marking and no dative alternation, we should expect that IO-DO would also be the preferred word order amongst Croatian children, and they might choose to produce it even when the givenness context is set up against it.

3. Research Questions and Predictions

The purpose of this study is to reveal which object order is the underlying one. By neutralising givenness and animacy, we can also establish whether these factors are triggers for movement. The study will also provide insight into which object order adult speakers use in set conditions. We will be able to compare adults and children, and thus observe whether they have the same tendency in neutral conditions and whether the two factors affect object order to the same extent.

The research questions are the following:

1. Do adults have a DO-IO preference in production as well?
2. What is the underlying object order in child language?
3. Are givenness and animacy triggers for movement in child language?
4. Do the triggers have an equal effect in the two groups of speakers?

In relation to our first research question, adults were found to have a preference for DO-IO when givenness and animacy were neutral. In the study reported on here, the role of adults is mainly as a comparison group for the children. In comparison to the acceptability judgment task from Velnić (2019), the results will provide insight on whether their preferences are the same in a production task. We thus expect a majority of DO-IO orders when animacy is balanced; nevertheless, the various givenness conditions are still expected to play a role in this setting.

With regard to the second research question, in a number of previous studies (section 2.3) a preference for the IO-DO order was noticed in children's productions. The languages that this was noticed for were case-marked, as

Table 1. Overview of the findings from this section

Study	Language	Dative Alternation	Case Marking	Age Range	Task	Givenness Effect on Word Order	Object Order Preference
Anderssen et al. (2014)	Norwegian	Yes	No	4–6	Production	Yes	PD (DO-IO)
Stephens (2015)	English	Yes	No	3;10–5;4	Production	Yes	PD (DO-IO)
Snyder (2003)	English	Yes	No	3;3–8;1	Corpus	Yes, increasing with time	NA
Höhle et al. (2014)	German	No	Yes	4;5–5;6	Production	Yes, but weaker than the word order effect[7]	IO-DO
Mykhaylyk Rodina, and Anderssen (2013)	Russian and Ukrainian	No	Yes	3–6	Production	Weak	IO-DO
Velnić (forthcoming)	Croatian	No	Yes	0;10–3;2	Corpus	No	IO-DO preference

[7] The sentences that violated the definiteness order were reproduced faithfully more frequently than the sentences that violated the DAT(IO)-ACC(DO) order, so there was an effect of givenness, but children were more likely to keep the preferred word order than to violate the definiteness order.

is Croatian, and as a result we might expect the same outcome in our task. However, as animacy was not balanced in the aforementioned studies but it is in a subset of our data, we mainly expect a majority of IO-DO productions when the IO is animate; whereas when both objects are animate we expect the children to be more adult-like.

Our prediction for the third question relates to the literature in sections 2.2 and 2.3. Since children were found to be very attentive to animacy in previous studies, we expect this to be a strong factor for movement. Givenness on the other hand might not be as strong due to the concept of non-shared assumptions (Schaeffer and Matthewson 2005) described in section 2.3.

Finally, the predictions for the last two research questions are intertwined. In light of previous discussions we expect children to be more attentive to animacy than adults (Snyder 2003) and their grasp of givenness not to be adult-like yet (Schaeffer and Matthewson 2005). In terms of object order, this will result in children producing more IO-DO orders in conditions of prototypical animacy. However, in relation to the condition where both objects are animate, we are unable to make any sound predictions. If the preference for IO-DO holds, then children should still produce a majority of IO-DO also when the DO is animate. Conversely, if they are aware of the underlying status of DO-IO, they will produce DO-IO more often, compared to the prototypical condition. The former outcome would confirm the preference for the IO-DO order, while the latter would be in favour of the high status of animacy as a trigger for movement. We expect adults to be more attentive to givenness throughout the task.

4. Methodology

In this section we outline the setup of the task used in our study.

4.1. Design

Our experiment tests two conditions of animacy and four conditions of givenness in order to check the effect of these factors as well as their interaction resulting in different object orders.

As mentioned before, I refer to the two animacy conditions as prototypical animacy (IO-animate and DO-inanimate) and balanced animacy (both animate). Animacy is set up as a binary feature, animate/inanimate: the referents of the task were either anthropomorphic animals or inanimate objects (e.g., a cat or an apple).

The four possible givenness conditions are the following: none of the characters are given (No-G); the DO is given (DO-G); the IO is given (IO-G); or all arguments are given (All-G). A referent is considered given if it has

been mentioned in the discourse. Thus in any first image of an experimental set, nothing is given, because none of the referents had the opportunity to be mentioned before. Following that, if the DO from the previous image is present again, this creates the DO-G condition; if the IO from a previous image is repeated, we have the IO-G condition. The conditions were each illustrated by one action image, with the exception of the No-G condition, which consisted of two images: one in which no argument was given and another in which the subject was given. They were merged under the No-G condition because in both of these conditions neither object is given and the givenness of the subject is not relevant for the current study.

This experimental design was inspired by the puzzle task developed by Eisenbeiss (2011) for eliciting a broad range of case-marked forms including double objects in German. Eisenbeiss's (2011) method consisted of a puzzle board with cut-outs containing images depicting various actions and puzzle pieces with the corresponding pictures to be put in the cut-outs. The children had to ask for the puzzle pieces corresponding to the pictures on the board and, since the pictures contrasted minimally one form the other, they were encouraged to mention all of the characters present in each picture. This method has proven to be successful, as it was engaging for the child and target structures were easily obtained. In order to control for givenness and animacy, we hereby adapt the method by setting up the conditions mentioned in the previous paragraph. The main difference from the original task is that here the participants begin with an empty puzzle board, and the images are provided by the experimenter.

The task consisted of a repeated-measures design, as for each value of animacy there is a variation of the four givenness types (2 x 4 = 8), and the aim of the task is to observe the interaction of the two factors. This was obtained through different sets of images, each one aiming to elicit a different verb. The set had either prototypical animacy (verb = *give*, *offer*) or balanced animacy (verb = *send*). Each set contained all givenness conditions presented, in the order as specified above (1. No-G; 2. DO-G; 3. IO-G; 4. All-G). The sets depicted a ditransitive action with the verbs *dati* 'give', *nuditi* 'offer', and *slati* 'send', respectively. The rationale behind the choice of verbs is that the verb 'give' is the most frequent ditransitive verb both in adult and child language (Kovačević 2004; Velnić 2014); the verb 'offer' was chosen because it can yield structural dative alternation of case (Zovko-Dinković 2007),[8] as briefly described in section 2.1. Lastly, 'send' was chosen in order to allow for balanced animacy, since it can accommodate an animate DO. Thus, the factor "verb" is not a variable of the design but merely a factor that allows us to set the animacy conditions of prototypical and balanced. By including also the possibility of dative alter-

[8] This alternation was not present in the children and had only two instances in the adult data.

nation, we have unfortunately rendered our design unbalanced, since in this configuration we have two sets of images with prototypical animacy and only one where animacy is balanced. An outline of the conditions is displayed in Table 2.

Table 2. Overview of the conditions in the task

Given DO	Given IO	Animate DO[9]
–	–	–
+	–	–
–	+	–
+	+	–
–	–	+
+	–	+
–	+	+
+	+	+

Object order is the dependent variable of the design, as the responses were labelled and analysed based on the object orders produced in the respective conditions. We will discuss the findings regarding our research questions based on this result.

4.2. Participants

A total of 59 monolingual Croatian children between the ages of 3;7 and 5;2 (mean age = 4;4, 26 males) were included in the task. We chose this age range because it is similar to the range used in previous studies that tested ditransitives (Anderssen et al. 2014; Höhle et al. 2014; Mykhaylyk Rodina, and Anderssen 2013; Stephens 2015). The children were recruited from four kindergartens in Rijeka, all part of a larger kindergarten group under the same administration. The parents had to sign an informed consent form in order for the children to participate.

The adult group functions as a background comparison group. It consisted of 36 participants aged 19–28 (mean age = 21, 8 males). The participants were required to have been born to Croatian speaking parents and to have grown up in Croatia; other languages learned later in life were not controlled for. They each received a 100 Kuna (approximately 13 euros) gift certificate at

[9] The IO was always animate.

a local bookstore for their participation. The participants were recruited at the Psychology and Law department of the University of Rijeka.

4.3. Materials

The materials for the experiment consist of the images depicting ditransitive actions (action images), images of single characters that are meant to fortify the givenness effect (single images), and the image board. All the images were printed on white Plexiglas. An example of the images is depicted in figures 1 below and 2 on the following page.

The action images depicted actions of transfer, and were divided into the three sets as already mentioned. Each set (n = 3) contained five action images (total = 15),[10] one for each givenness condition. The images were shaped differently from one another, and each set had one image corresponding to one shape on the board. We have also controlled for directionality: the order in which the referents (e.g., the agent and the recipient) are drawn varies (either left to right, or right to left), with the DO always placed in the middle in order to provide a clear depiction of the referents' interactions.

The single images depicted one of the referents present in the action images. Their role was to reinforce the givenness condition, as they were presented in-between action images and contained a referent present in the

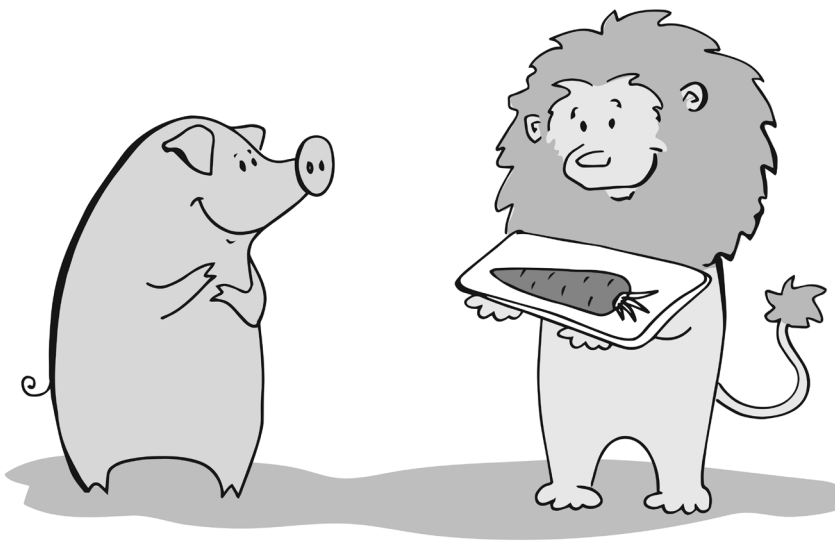

Figure 1. Action image (from the 'offer' set)

[10] Recall from section 4.1 that the No-Given condition consisted of two images.

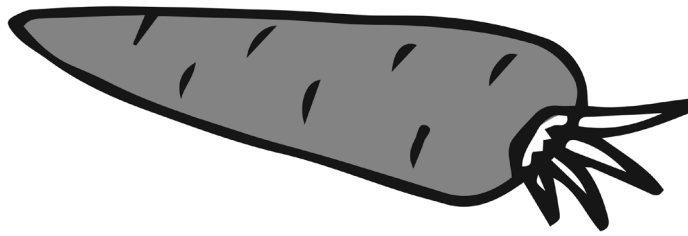

Figure 2. Single image (from the 'offer' set)

previous and in the following action image. Each set contained four single images (total = 12).

All these images had to be placed on the board. The image board consisted of two wood planks attached to one another, with the top one containing five differently-shaped slots, one for each action image. At the bottom of the board there was a small shelf designated for the single images (subject, theme, or recipient) that reinforce which one is given in the following action image. An example of the board with some images placed on it is provided in figure 3 below.

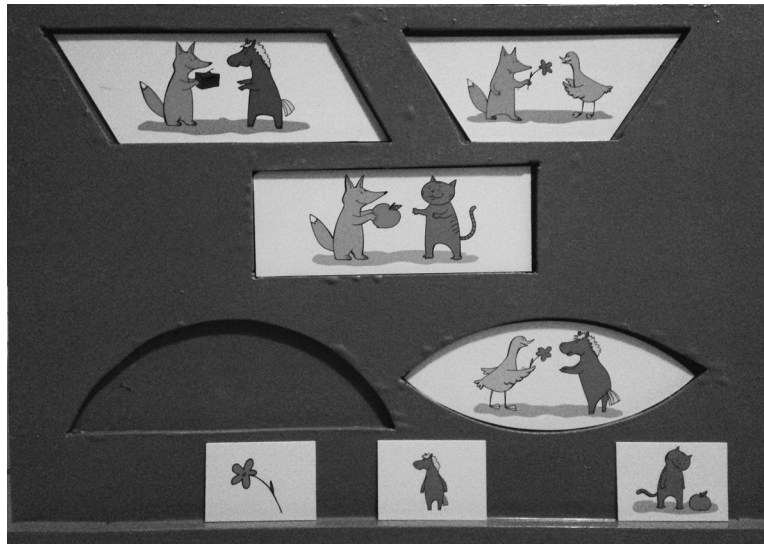

Figure 3. Photograph of the image board with some images on it.[11]

[11] NB: The way the images are placed on the board in the photograph does not exemplify a real situation in the experiment

4.4. Procedure

All the sessions were audio recorded. The recordings took place in a room on the kindergarten premises, where the child and the researcher could remain undisturbed. For the adults, the testing was conducted either in the psychology lab or in a classroom at the university of Rijeka. An audio recorder (model: Sony lcd-px333) was placed on the table, and the experimenter also manually recorded the children's responses as the testing proceeded. This was then used to facilitate the transcription process. The responses were not manually recorded for the adult controls because the testing proceeded very smoothly, and the on-line transcription would have slowed down the task.

The distribution of the previously-mentioned shapes was different for each set. The shapes are not relevant for the study; their function was to make the task more entertaining for the child and also to add more cognitive load to the task so there is less chance for auto-priming. The images had the same order of givenness conditions across the sets: No-G, DO-G, IO-G, and All-G. There were two possible orders in which the images of a set could be presented, but the order of the givenness conditions remained unvaried. One of the orders in which the images were presented to the participants is shown in Tables 3–5 for each verb. The referents (animals and objects) are different in every set so as to avoid cross-condition givenness effects. Note that the descriptions in the tables below are merely describing what is drawn on the action image and do not reflect our expectations or the actual productions of the participants.

Table 3. One possible order of images for 'give'

	Given	Action	Direction
1	**No given**	Fox gives apple to cat.	S > DO > IO
2	**No given**	**Fox** gives flower to duck.[12]	S > DO > IO[13]
3	**S & DO**	**Duck** gives **flower** to horse.	S > DO > IO
4	**S & IO**	**Fox** gives cake to **horse.**	S > DO > IO
5	**All**	**Duck** gives **apple** to **cat.**	IO < DO < S

[12] In this image the subject is given, but, as in the former condition neither object is given, and they are thus counted under the same condition. The layout is the same for all the sets.

[13] This image was originally supposed to have the IO < DO < S order, and it was illustrated that way, but during the printing process it was reversed and printed as a mirror image, which resulted in the inverse orders of the characters.

Table 4. One possible order of images for 'offer'

	Given	Action	Direction
1	**No given**	Lion offers lollipop to zebra.	S > DO > IO
2	**No given**	**Lion** offers carrot to pig.	IO < DO < S
3	**S & DO**	**Pig** offers **carrot** to monkey.	IO < DO < S
4	**S & IO**	**Lion** offers sandwich to **monkey.**	S > DO > IO
5	**All**	**Pig** offers **lollipop** to **zebra.**	IO < DO < S

Table 5. One possible order of images for 'send'

	Given	Action	Direction
1	**No given**	Bunny sends puppy to elephant.	IO < DO < S
2	**No given**	**Bunny** sends parrot to turtle.	S > DO > IO
3	**S & DO**	**Turtle** sends **parrot** to snail.	IO < DO < S
4	**S & IO**	**Bunny** sends mouse to **snail.**	IO < DO < S
5	**All**	**Turtle** sends **puppy** to **elephant.**	S > DO > IO

The second order in which the images could be presented to a participant is provided in the appendix. Thirty-four of the children received the images in order 1 (presented in Tables 3–5), while 24 were presented with order 2. This imbalance is due to the fact that the two orders of images were presented on alternating days and on some days, there were more children tested than on other days. In the control group, 18 participants were given the images in order 1 and 18 in order 2. Figure 4 on page 102 illustrates the task of presenting the images to the participants according to order 1 of 'give'.

The task proceeded as follows. The experimenter and the participant sat opposite each other. The image board was located in front of the participant, positioned in such a way that the experimenter could not see what was being placed on it. The participant was instructed to receive the images, describe them, and place them in the appropriately-shaped slot. At the beginning of each puzzle set, the experimenter prompted the verb by saying "these images are about giving/sending/offering". The sets were given in a random order.

The images were given to the participant from a bag, facing down, so that the participant was the only one to see the image. The action images depicted the actions regarding the respective verb and involved three referents: the agent, the theme, and the recipient. They were thus targeted to elicit a ditransitive structure, which appropriately describes the interaction of the referents.

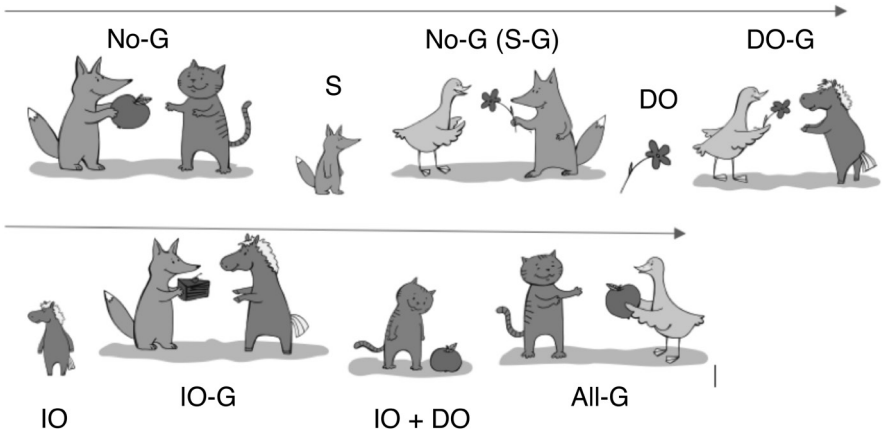

Figure 4. Order 1, in which the images were given to the participant to elicit 'give'.

After each action image the participant received a single image of the referent that was present in the previous image, and that was about to also appear in the next action image. The experimenter and participant exchanged a few sentences about it before proceeding to the next action image. The conversation usually consisted of the experimenter asking the participant whether this referent was the same one as seen in the action image or asking the participant whether they liked the referent on the single image. The latter strategy was more successful with children than with adults; the adults were not keen on expressing their liking for a referent. This was repeated until all five images of a set had been described and placed on the board. Once the board was complete, the experimenter and the participant took out all the images, the board was placed in front of the participant once more, and they proceeded with the next set of images. This was repeated for all three verb sets. At the end of the task the child was accompanied back to the kindergarten group, while the adult was given the reward.

As noted above, the sets were supposed to be given in a random order. However, after a few runs we noticed that the 'send' set had less data loss in the children's productions if presented last. This set was harder than the other two, most likely due to an unprototypical situation of sending an animate referent to another animate referent. By having this set as the last one, the child was familiar with the procedure and thus described the images more easily. We therefore proceeded by randomly giving one of the two IO-animate sets as first and second, while the both-animate set was given last.

5. Results

The results will be reported separately for children and adults, and the two groups will be compared in the discussion. We start by accounting for the non-applicable data and then we will continue by describing the results with their respective statistical analyses, separately for the two groups. The statistical analysis includes linear mixed effects (Bates et al. 2015), used to establish the various models: a null model with no special attention to either factor and a separate model focusing on animacy and givenness respectively. ANOVAs were used to establish the significance of a factor model with respect to the null model and also to observe the potential interaction of the two factors. The analysis then proceeds with a pairwise comparison of the givenness conditions, separately for the two animacy conditions. This way of approaching the data provides us with an in-depth understanding of how animacy and givenness affect word order: the effect of the individual factors, their interaction, and how each condition shapes word order. Additionally, for the child data we have set up a linear mixed effect on the full data set in order to observe the effects of the two factors more thoroughly.

5.1. Non-Applicable Data: Production Exclusions

The adult controls had 540 possible responses (5-targets x 36-adults x 3-sets), and we were able to use 439 of them. The NA data were due to the following: no ditransitive action (n = 19), inverted referents (n = 6), the use of clitics (n = 6), and the use of a PP (n = 70), which was excluded due to end-weight affecting the object order. An example of each of these NA responses is given in the examples below:

(4) No ditransitive action:

Zec tjera neku drugu životinju a
rabbit$_{NOM}$ chase$_{PRES.3SG}$ some$_{ACC}$ other$_{ACC}$ animal$_{ACC}$ and
kornjača to gleda.
turtle$_{NOM}$ that watch$_{PRES.3SG}$

'The rabbit is sending away some other animal, while the turtle is watching.'

(5) Inverted referents:

Kornjača pokazuje slona psu.
turtle$_{NOM}$ show$_{PRES.3SG}$ elephant$_{ACC}$ dog$_{DAT}$

'The turtle is showing the elephant to the dog.'
Target: The image depicted the turtle sending the dog to the elephant.

(6) Use of clitics:

Ovdje je slon i zec mu šalje
here is elephant$_{NOM}$ and rabbit$_{NOM}$ him$_{CL.DAT}$ send$_{PRES.3SG}$
psića još jednog.
doggy$_{ACC}$ more one$_{ACC}$
'Here is the elephant and the rabbit is sending him another doggy.'

(7) Use of a PP:

Zec šalje pticu kod kornjače.
rabbit$_{NOM}$ send$_{PRES.3SG}$ bird$_{ACC}$ at turtle$_{GEN}$
'The rabbit is sending the bird to the turtle.'

The children strongly overused the verb 'give' across all conditions, which still yielded a ditransitive. We are not excluding these data, as we were not interested in testing the word order with a particular lexical verb but in the effect of animacy and givenness on object order combinations. Out of 885 responses (5-targets x 59-children x 3-sets), we were able to use 625. The NA child data are categorized as follows: no response (n = 5), no ditransitive action (n = 74), use of subordinate clause (n = 39), case error with non-intelligible roles (n = 6), referent inversion (n = 67), omission of an object (n = 58), use of a pronoun or clitic (n = 10), and experimenter's mistake (n = 1). An example of no ditransitive action and case error with non-intelligible roles is provided below, as these are straightforward and will not be discussed any further, whereas the other examples will be provided along with an explanation of the error.

(8) Lav uzme lizajku onda donese kući.
lion$_{NOM}$ take$_{PRES.3SG}$ lollipop$_{ACC}$ then bring$_{PRES.3SG}$ home
'The lion took the lollipop and brought it home.'

(9) Lav daje mrkvu svinju
lion$_{NOM}$ give$_{PRES.3SG}$ carrot$_{ACC}$ pig$_{ACC}$

In (9), since both objects are given in the accusative, there is no morphosyntactic way of telling the roles of theme and recipient apart. The roles could be disambiguated through animacy, and we can assume that the pig is meant to be the recipient, but we nevertheless decided to exclude examples like these. As can be seen from the very limited number of these errors, the children we tested had acquired case and had no problem marking the two objects distinctly with the appropriate morphology.

Since PPs and subordinate clauses are more likely to be heavy and thus be placed at the end of the sentence, we have decided to exclude them from

the dataset being analyzed. An example of a sentence with both a subordinate clause and a PP is displayed in (10).

(10) Šalje zec da tamo ode kod slona.
send$_{PRES.3SG}$ bunny$_{NOM}$ that there go$_{OPT.3SG}$ at elephant$_{GEN}$
'The bunny is sending it (the dog) to go there to the elephant.'

The referent inversion occurs when the child inverts the IO and DO roles, by assigning the dative case to the target DO and the accusative to the target IO. This was not a case mistake, since the children use the cases correctly in the other sets. Even though we have accepted deviations from the intended verb, the inversion of the theme and recipient is a description of a different event entirely, and also influences the givenness conditions. All of the referent inversions were confined to the both-animate condition, where it was possible to invert the DO and the IO. An example, along with the target description is given in (11).

(11) Ovdje zec pokazuje mišiću puža.
here rabbit$_{NOM}$ show$_{PRES.3SG}$ mouse$_{DAT}$ snail$_{ACC}$
'Here the rabbit is showing the mouse the snail.'
Target: The image depicted the rabbit sending the mouse to the snail.

Even though the use of a pronoun or a clitic is an indication of givenness, we have decided to exclude these forms, because they also influence word order, as a pronoun is usually placed before an NP, while clitics are syntactically fixed in second position. An example of the use of the clitic is provided in (12).

(12) Konj mu je dao cvijet.
horse$_{NOM}$ him$_{CL.DAT}$ AUX gave flower$_{ACC}$
'The horse gave him a flower.'

The children's object omissions will be discussed separately, in section 5.5.

5.2. Intended Givenness vs. Actual Givenness

During the test, the child would often take an image, say what was on it, and then describe the action. In such cases, all the referents have to be counted as given. This problem only occurred infrequently in adults, as they typically did not mention anything prior to the ditransitive target.

A crucial part of the data analysis is to observe how word order changes in relation to givenness. We thus had to account for what was actually given and re-categorize the occurrences accordingly. Table 6 shows the final count

of responses in each condition. Although adults did not deviate from the intended givenness condition, their number of responses is nevertheless provided in Table 7. This is the final distribution of the data that will be analysed and discussed in the next section.

Table 6. Distribution of responses in the actual state of givenness in the child data

Condition	**No-G**[14]	**DO-G**	**IO-G**	**All-G**
N. responses	180	127	149	169
Total		625		

Table 7. Distribution of responses in the adult data

Condition	**No-G**	**DO-G**	**IO-G**	**All-G**
N. responses	177	86	91	85
Total		439		

5.3. Adults' Responses

We will first outline the responses obtained by the adult participants. From a look at the raw data it is evident that adults produce more DO-IO orders, especially in the presence of balanced animacy. Some examples follow.

(13) DO-G Prototypical animacy

a. Patka daje cvijet konju.
duck$_{NOM}$ give$_{PRES.3SG}$ flower$_{ACC}$ horse$_{DAT}$
'The duck is giving the flower to the horse.'

[14] Recall that the No-G condition includes two images for each set: No-G and Subject-G, because neither object is given in both of those conditions.

(13) b. Svinja nudi mrkvu majmunu.
pig$_{NOM}$ offer$_{PRES.3SG}$ carrot$_{ACC}$ monkey$_{DAT}$
'The pig is offering the carrot to the monkey.'

(14) IO-G Prototypical animacy

a. Lisica daje konju čokoladnu tortu.
fox$_{NOM}$ give$_{PRES.3SG}$ horse$_{DAT}$ chocolate cake$_{ACC}$
'The fox is giving the horse some chocolate cake.'

b. Lav nudi kekse majmunu.
lion$_{NOM}$ offer$_{PRES.3SG}$ cookies$_{ACC}$ monkey$_{DAT}$
'The lion is offering some cookies to the monkey.'

(15) DO-G Balanced animacy

Tu kornjača šalje papigu pužu.
here turtle$_{NOM}$ send$_{PERS.3SG}$ parrot$_{ACC}$ snail$_{DAT}$
'Here the turtle is sending the parrot to the snail.'

(16) IO-G Balanced animacy

Zec šalje miša pužu.
bunny$_{NOM}$ send$_{PRES.3SG}$ mouse$_{ACC}$ snail$_{DAT}$
'The bunny is sending the mouse to the snail.'

The general reasoning behind the choice of statistical analysis has been outlined in section 5. Tables 8–10 provide a summary of the ANOVAs of the null model with the animacy and givenness model respectively, along with the ANOVA conducted on the interaction.

Table 8. ANOVA of the null and animacy model in the adult data

Model	Df	AIC	BIC	Chisq	Significance
Null	2	469.40	477.59	53.416	$p < 0.001$
Animacy	3	417.98	430.28		

Table 9. ANOVA of the null and givenness model in the adult data

Model	Df	AIC	BIC	Chisq	Significance
Null	2	469.4	477.59	14.193	p < 0.01
Givenness	5	461.2	481.69		

Table 10. ANOVA of the interaction of animacy and givenness in the adult data

	Df	AIC	BIC	Chisq	Significance
Interaction	6	406.55	431.14	1.821	No
	9	410.73	447.61		

The data clearly show that both factors have an effect on word order, but animacy is stronger. This is evident from the fact that there is no interaction, and from the depiction of the data displayed in figure 5 on the following page. Thus, animacy shapes object order, and the effect of givenness influences the object order within the animacy condition. A possible reason for this may be the design, as there was only one image set with balanced animacy. Another reason might be data loss from that set causing the two animacy conditions to have an unbalanced number of observations. This is of course the limit of our task, but it is nevertheless obvious that animacy is a stronger factor than givenness in affecting object order choice.

The next step is to look into the effect of the individual conditions (total = 8), and we will do so by conducting a pairwise comparison of the givenness conditions, separated into two animacy conditions. The results of the statistical analysis are displayed in Tables 11 below and 12 on the following page.

Table 11. Pairwise comparison of the givenness conditions in adults when animacy is prototypical

Contrast	Estimate	SE	Z-ratio	Significance
ALL-NO	0.335	0.369	0.910	No
ALL-DO	-1.238	0.462	-2.679	p > 0.05
ALL-IO	0.120	0.421	0.286	No
NO-DO	-1.574	0.408	-3.859	p < 0.001
NO-IO	-0.215	0.356	-0.505	No
DO-IO	1.359	0.454	2.991	p > 0.05

Table 12. Pairwise comparison of the givenness conditions in adults when animacy is balanced

Contrast	Estimate	SE	Z-ratio	Significance
ALL-NO	-0.848	1.507	-0.563	No
ALL-DO	-16.753	111.757	-0.150	No
ALL-IO	-16.614	106.787	-0.156	No
NO-DO	-15.905	111.771	-0.142	No
NO-IO	-15.766	106.794	-0.148	No
DO-IO	0.138	156.321	0.001	No

The statistical results presented in Table 11 clearly show that the condition with the given DO stands out, as it is significantly different when compared to the three remaining conditions. Figure 5 also highlights the DO-given condition, as the participants produce significantly more DO-IO orders than in the rest of the task. Conversely, when animacy was balanced there is no difference between the conditions. This is due to the fact that DO-IO is used at ceiling level, and consequently there is no variation in the responses—a fact evident from the negative values of the results, which signal a DO-IO preference over IO-DO.

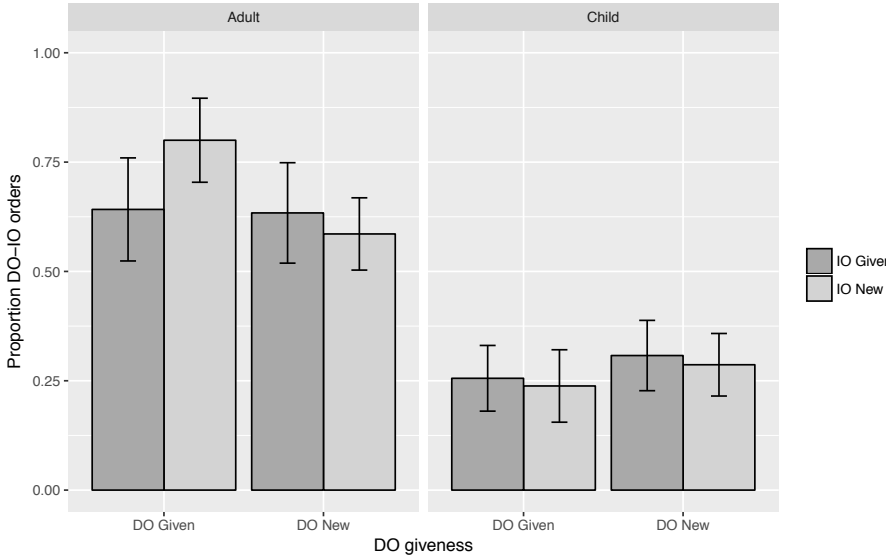

Figure 5. Proportion of DO-IO orders used by adults in the task

Figure 5 clearly shows a great difference of object order distribution between the two animacy conditions; it also surprisingly shows that DO-IO productions are at ceiling level when both objects are animate. This might, however, be a task effect, which will be elaborated on in the discussion. The givenness effect of the DO-given is also evident for the prototypical animacy condition, given that we can clearly see how the production of the DO-IO is increased when compared to the other givenness conditions. Overall, the adults produce more DO-IO occurrences than IO-DO, even when only the IO is animate. This means that even if the animacy effect has been found to be strong (based on the comparison of object order productions in the two animacy conditions), there is still the preference to have the DO precede the IO, even when only the latter is animate.

5.4. Children's Results

In contrast to the adults children have a preference for IO-DO when animacy is prototypical, but the two word orders are at chance level when it is balanced. Some examples follow.

(17) DO-G prototypical animacy

a. A onda prase daje majmunu mrkvu.
and then pig$_{NOM}$ give$_{PRES.3SG}$ monkey$_{DAT}$ carrot$_{ACC}$
'And then the pig is giving the monkey the carrot.'

b. Patka je dala cvijet konju.
duck$_{NOM}$ AUX gave flower$_{ACC}$ horse$_{DAT}$
'The duck gave the flower to the horse.'

(18) IO-G Prototypical animacy

a. Tu je lav i pokazuje majmunčiću keksiće.
here AUX lion$_{NOM}$ and show$_{PRES.3SG}$ monkey$_{DAT}$ cookies$_{ACC}$
'Here is the lion and he is showing the monkey some cookies.'

b. Lisica dava konju tortu.[15]
fox$_{NOM}$ give$_{PRES.3SG}$ horse$_{DAT}$ cake$_{ACC}$
'The fox is giving the horse a cake.'

c. Vjeverica je dala kolač konju.
squirrel$_{NOM}$ AUX give$_{PST.3SG}$ cake$_{ACC}$ horse$_{DAT}$
'The squirrel gave a cake to the horse.'

[15] *Dava* is an inflection mistake, the correct form would be *daje* as used in the other examples. Nevertheless, this kind of error does not affect our results.

(19) DO-G Balanced animacy

a. Tu je kornjača dala papigu pužu.
here AUX turtle$_{NOM}$ give$_{PST.3SG}$ parrot$_{ACC}$ snail$_{DAT}$
'Here the turtle gave the parrot to the snail.'

b. A kornjača pužu daje pticu.
and turtle$_{NOM}$ snail$_{DAT}$ give$_{PRES.3SG}$ bird$_{ACC}$
'And the turtle is giving a snail the bird.'

(20) IO-G Balanced animacy

a. Zec dava pužu miša.
bunny$_{NOM}$ give$_{PRES.3SG}$ snail$_{DAT}$ mouse$_{ACC}$
'The bunny is giving the snail a mouse.'

b. Zec želi dati miša pužu.
bunny$_{NOM}$ want$_{PRES.3SG}$ give$_{INF}$ mouse$_{ACC}$ snail$_{DAY}$
'The bunny wants to give a mouse to the snail.'

The same setup has been used for the child data as well. The ANOVAs of the null model with the animacy and givenness model, respectively, are reported in Tables 13 and 14, along with the interaction in Table 15 on the following page.

Table 13. ANOVA of the null and animacy model in the child data

Model	Df	AIC	BIC	Chisq	Significance
Null	2	701.01	709.92	33.421	$p < 0.001$
Animacy	3	669.59	682.95		

Table 14. ANOVA of the null and givenness model in the child data

Model	Df	AIC	BIC	Chisq	Significance
Null	2	701.01	709.92	1.9959	No
Givenness	5	705.01	727.28		

Table 15. ANOVA of the interaction of animacy and givenness in the child data

	Df	AIC	BIC	Chisq	Significance
Interaction	6	672.43	699.15	2.9151	No
	9	675.51	715.60		

From the data in these tables we can see that animacy affects word order very strongly. However, there is no effect of givenness on word order in the child data. As in the adult data, an interaction between animacy and givenness has not been found.

Thus here we have already found a relevant difference between children and adults, as children seem to be unaffected by givenness in realizing object order. This also fits with previous findings and with our predictions: animacy was claimed to be easily acquired, while opinions were divided regarding givenness and its visibility in the effect of word order.

A pairwise comparison was also conducted with the child data for each givenness condition, separately for when animacy is balanced and prototypical. No significant differences were found between any two conditions. This is to be expected, as we have found no effect of givenness with the previous test. The tables containing the full results of the pairwise comparison can be found in the appendix. The distribution of DO-IO orders per condition is depicted in Figure 6 on the following page.

It is strikingly evident that children use DO-IO to a much lesser degree than the adults overall, but it is also evident that the proportion of DO-IO increases when animacy is balanced. The animacy effect we observed is shown in Table 13. The statistical analysis found no effect of givenness, and we can see that within each animacy condition the distribution of DO-IO is roughly the same and does not change based on what is given. The givenness effect was isolated to the DO-given condition within the adult data. If we take a closer look at the distribution of DO-IO within this condition, we can see that there is a decrease in the number of occurrences of this order, which indicates a new > given preference. However, when animacy is balanced, the DO-G condition has an increased production of DO-IO, as we would expect. We have thus set up a model using linear mixed effects consisting of the two animacy conditions taking the givenness conditions into account (Table 16 on page 113). The givenness contrasts were set up based on the givenness of the DO: DO-G and All-G on the one hand and No-G and IO-G on the other were grouped together in groups, which are tagged as DO-GG and DO-nG respectively. The intercept is the children's responses in the prototypical animacy condition.

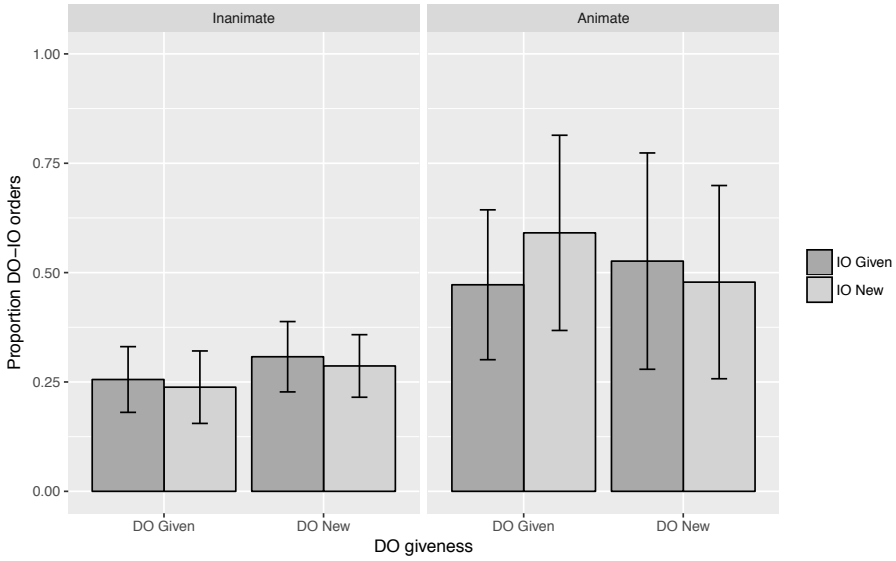

Figure 6. Proportion of DO-IO orders used by children in the task

The significance in the intercept indicates a preference for one object order in the IO-animate condition. This preferred object order is IO-DO, as the estimate has a (–) sign. Furthermore, we can see that the production of DO-IO significantly decreases when we compare the conditions with a given DO to

Table 16. Summary of the model for children's responses in the two animacy conditions, with relation to givenness

	Estimate	Standard error	p-value	Significance
IO-animate (Intercept)	-1.4135	0.26244	7.20e-08	$p < 0.001$
IO-animate DO-GG vs DO-nG	-0.9289	0.4725	0.0493	$p < 0.05$
IO-animate DO-G vs All-G	-0.2025	0.3531	0.5663	No
IO-animate IO-G vs No-G	0.3472	0.3103	0.2632	No
Both-animate	1.6436	0.3106	1.22e-07	$p < 0.001$
Both-animate DO-GG vs DO-nG	2.0227	1.1289	0.0732	$p < 0.1$
Both-animate DO-G vs All-G	0.6487	0.773	0.4013	No
Both-animate IO-G vs No-G	0.0057	0.8124	0.9943	No

the conditions where DO is not given ($p < 0.05$). The givenness of the IO does not seem to be of any relevance, as the two subsequent comparisons do not come out as significant. The comparison of the data in the two animacy conditions reveals that children use significantly more DO-IO when both objects are animate. Moreover, the comparison of the DO-GG and DO-nG is almost significant ($p < 0.1$), entailing that the DO-IO increases in conditions of given DO, contrary to what happens when only the IO is animate. The reason why this interaction is not significant might be the reduced amount of data elicited for the both-animate condition. Thus if the conditions had been comprised of an equal amount of sets, the result of this interaction would most likely have been significant. The givenness of the IO does not seem to play a role.

5.5. Omissions in the Child Data

Previous studies, such as Mykhaylyk, Rodina, and Anderssen (2013) and Anderssen et al. (2014), found a significant amount of data related to givenness in the omissions. Since the production of object order does not signal sensitivity to givenness in the child data, we decided to check if the omissions are related to it.

Overall, the children have 58 object omissions, 42 of omitted elements being given. The adults did not have any omissions in the task. Table 17 on the following page shows this omission by element across the givenness conditions; the shaded values signal that the argument is given. Some of the omission examples are displayed in (21).

(21) a. Ovaj daje čokoladni kolač.
this$_{NOM}$ give$_{PRES.3SG}$ chocolate cake$_{ACC}$
'This one is giving a chocolate cake.'

b. Prasac je dao majmunu.
pig$_{NOM}$ AUX give$_{PST.3SG}$ monkey$_{DAT}$
'The pig gave the monkey.'

Most omissions occur in the All-G condition, and the IO has the highest omission rate (n = 44). The most relevant omissions are DO and IO omissions in the DO-G and IO-G conditions, as these can signal whether the omission is related to givenness. Table 18 on the following page shows the distribution of these omissions along with the occurrences containing both objects. The shaded values signal an appropriate construction or omission in relation to givenness.

Table 17. Distribution of omissions in the child data

	No-G	DO-G	IO-G	All-G	Total
DO	1	2	2	9	14
IO	9	4	10	21	44
Total (omitted+overt)	154	108	126	171	559

Table 18. Distribution of word orders and omissions in DO-G and IO-G

	IO-animate		Both animate	
	DO-G	IO-G	DO-G	IO-G
DO-IO	23	38	12	10
IO-DO	60	56	4	5
Om DO	2	2	0	0
Om IO	1	7	3	3
Total appropriate productions	25	63	12	8

We can see that the omissions are marginal in the key conditions for this study, and we can make very few observations on the omission pattern. Firstly, the IO is much more prone to omission than the DO. Overall, children omit slightly more given objects than new objects (12 vs. 6). However, these data are too scarce to suggest that children mark givenness through the omission of the given object rather than through word order, as both strategies (the IO-DO order and the omission of the IO) show non-context-related preferences.

6. Discussion

The task in our study was to reveal the underlying order in ditransitives in child language and to explore whether animacy and givenness were triggers for movement and how that compares to the adults' productions.

Our first question was whether DO-IO surfaced as the underlying order for adults in this production task, as it has been found in an acceptability judgment task (Velnić 2019). And indeed it did, as we found DO-IO being produced at ceiling level when animacy was balanced. The ceiling level was not the expected result, but it points quite strongly to the status of DO-IO as underlying. When compared to the acceptability judgment task, animacy seems to trigger a majority of DO-IO in both settings. This result provides cumulative evidence to the body of research so far on Croatian ditransitives that

DO-IO is the underlying order. When animacy was prototypical, the DO-IO proportions were roughly at chance level, which is still quite different from what was found in the corpus data (Velnić forthcoming). A more thorough examination is needed to reveal whether the cause of this is the exclusive use of NPs to express the objects (compared to the corpus data, which contain pronouns and clitics). The givenness effect was confined to the condition of given DO, which indicates that the givenness of the IO is not relevant for ordering the two objects. More investigation is needed to find out why.

The production of DO-IO was at ceiling level when both objects were animate. A possible reason for finding a limited givenness effect in the adult controls is that the task may have failed to distinguish between given and new elements. Perhaps the adults did not believe that the experimenter did not know which images she was taking out of the bag. In that case, they might have perceived everything as given and thus did not have the need to mark givenness distinctly. Either way, the results of the statistical analyses have shown that both animacy and givenness have an effect on object order in the adult language.

The second research question was related to the underlying word order in the child data. Previous studies have found, as outlined in section 2.3, that children have a preference for IO-DO. This is of course true if we look at only prototypical animacy. Nevertheless even when animacy is prototypical, the children do not produce their preferred order at ceiling level but somewhat stably at around 70% of the time (all givenness conditions averaged). However, when animacy is balanced, the productions of DO-IO and IO-DO reach a chance level, and thus there is no longer a preference for IO-DO. It would seem that since the previous studies discussed in the literature review did not balance for animacy, the tendency to produce IO-DO was caused by the animacy of the IO. We can safely say that our prediction of a majority of IO-DO productions related to animacy has been borne out. However, children are not more adult-like when animacy is balanced, even though the increase of DO-IO is considerable. This is due to the fact that adults produced DO-IO at ceiling level, which was not an expected outcome. We have already suggested that this is probably a task effect, so it is likely that children are indeed more adult-like when animacy is balanced. Unfortunately this task is not able to show it.

This brings up the discussion regarding animacy and givenness as triggers for movement (research question 3). The fact that there is a significant difference in the proportion of object orders in the two animacy conditions means that animacy is a strong factor for ordering the two objects. But we found no general effect of givenness. This matches our predictions, as previous studies have shown time and time again how children are attentive to animacy, whereas their attentiveness to givenness was mixed. However, when the child data are analysed more closely, we can see that the givenness of the

DO is more relevant than the givenness of the IO. This was also found in the adult data but with a more prominent degree.

However, what does this mean in relation to the underlying word order in Croatian children? We have stated in the predictions that if children had a preference for IO-DO, this order should be produced more often regardless of the animacy or givenness condition. But if they were aware of the underlying status of the DO-IO in Croatian, there should be a significant increase of this order when animacy is balanced. Our results clearly point in the direction of the latter conclusion, even if the underlying word order is not as clearly available as for the adults (DO-IO), which could mean that children are not adult-like yet. In the balanced condition, the children were more adult-like, as they produced DO-IO 52% of the time. However, it is obvious that IO-DO is not the underlying order, because if it were it could be used at ceiling level in both conditions. Since we have established that givenness is not a trigger for movement, animacy favours IO-DO in the prototypical condition. But it is not a factor when balanced. Thus if children were guided by a combination of animacy and what for them is underlying, IO-DO could be predicted across the task. The fact that DO-IO is produced significantly more when animacy is balanced shows a tendency for children to prefer this order; in turn, it could potentially also indicate that they are sensitive to the same factors as the adults are, but not to the same extent.

Our last research question is about the comparison of adults and children. Animacy is obviously a trigger for movement in both groups, but to a greater extent in the children's group, as they use IO-DO significantly more than the adults when the IO is animate. The data also reveal that givenness has an effect on object order in the adult group. While the overall effect of givenness was not seen in the child data, we find that the givenness of the DO is more relevant than the givenness of the IO in both groups.

In relation to the underlying word order, it is obvious that it is DO-IO in the adult group. The children also show a tendency for this order, but it is not as pronounced as in the adults. What the data have definitely shown is that the children do not have a tendency for IO-DO, as reported for some other languages, but that their preference for IO-DO is limited to the condition of prototypical animacy, which is the one found in naturalistic data. With regard to our predictions, children seem to be more attentive to animacy than the adults, which is obvious from the proportions of IO-DO when the IO is animate. Children are also less attentive to givenness than adults, as we found no significant effect of givenness on object order, whereas the adults have an effect confined to the given DO. We have predicted that, if for the children the DO-IO is underlying, they should produce it more when animacy is balanced. And this is indeed the case. This outcome supports the high status of animacy as a trigger for movement.

The results suggest that animacy is a strong factor for determining word order in both types of speaker, more significantly in children. The results can also be discussed in light of the interaction of animacy with the different object order preferences in adults and children. More precisely, and in light of the other data on Croatian, such as the acceptability judgments obtained by Velnić (2019), it is obvious that the preferred word order of adult speakers is DO-IO. Therefore, since animacy influences object order choice, when the IO is animate the adults produce their preferred order and the animate-first order in equal proportions. The production of DO-IO is increased with givenness in favour of the DO (DO-G condition). It then returns to the initial distribution, which is the interaction of word order preference and the animacy of the IO. The givenness of the IO does not seem to be a factor.

When animacy is no longer a factor (the objects are both animate), adults produce DO-IO at ceiling level, as their word order preference is the only ordering mechanism that surfaces. The reason for this is open to discussion, since we expected adults to be the group that takes more factors (in our case givenness and animacy) into consideration when ordering the arguments. It nevertheless seems, contrary to any prediction, that adults choose based on the pragmatic availability of their preferred order, and that once free from animacy constraints they use that order exclusively. It is peculiar that givenness is completely ignored here, but we have already mentioned that this might be due to a task effect in which the adults considered all referents as given. If that is the case, animacy is the only factor tested on adults, and it has an effect that we have already discussed.

For the children, naturalistic data from Croatian suggests that IO-DO is the more frequently produced object order (as per the corpus and experimental studies discussed in sections 2.2 and 2.3). This is not strictly an indication of their preference for this order, since child-directed speech also contains a majority of IO-DO (Velnić 2014).

Let us, then, first outline the children's behavior in our task and see whether there really is a preference for IO-DO. In the IO-animate condition, children produced mostly IO-DO because it is the more appropriate object order from an animacy perspective, to which we know children are attentive. The production of DO-IO significantly increases when animacy is balanced, entailing that it is a very relevant factor. If IO-DO was really their preferred order, it could have been used unvaryingly across the task, since its use is still appropriate from an animacy perspective. Here, the children also show a givenness effect similar to that observed in adults for the prototypical animacy condition, as the DO-G condition has more DO-IO productions than the other givenness conditions. Perhaps, once animacy is balanced, children have more cognitive capacity to integrate givenness in their word order choice. This is only speculation, and there is no way of proving it based on the available data.

However, children do not reach ceiling level in any condition, as adults do in the balanced animacy condition. The data suggest that children do not prefer IO-DO and are aware of the underlying status of DO-IO, but are not yet adult-like. If they relied only on the appropriateness of IO-DO, they could have used it consistently throughout the task. Thus, the predominant productions of IO-DO seen in the naturalistic data and in some of the experimental studies cited here are due to the animacy imbalance and children being very sensitive to it. Once that is removed, children speakers are freer to vary their productions and be more similar to the adults.

To conclude, the object order choice we see in the task is the interaction of preferred object order and animacy. Animacy seems to have a stronger effect on children than on adults, which is in line with what Snyder (2003) had found.

7. Conclusion

This study set out to explore the underlying order of direct and indirect objects and the effects of givenness and animacy as triggers of movement in the ditransitive sentences of Croatian pre-school children.

Although we found a strong animacy effect in both groups of speakers, we concluded that children rely on animacy more than adults. An effect of givenness was found only in the adult group, and it was limited to the condition in which the DO was given. Children were also more attentive to the DO being given, but this was statistically marginal, as seen in the stable distribution of IO-DO in the condition of prototypical animacy. This was a predicted result for the children (but not for the adults), as we expected them to take the givenness of all the arguments into consideration. The reasons for why the givenness of the IO does not trigger an effect of word order are yet to be investigated.

We have also confirmed that adults prefer DO-IO in production as well, while children tend to use more IO-DO but do not have a strong preference for that object order. In the child data there is an over-production of IO-DO when the IO is animate, but once animacy is balanced the proportion of the two word orders is in favor of DO-IO. The predominance of IO-DO productions in naturalistic data is due to the IO being animate and the DO being inanimate. This study shows that once animacy is no longer a factor, the DO-IO preference surfaces. This suggests that children are very attentive to animacy but that their word-order preference is underlyingly adult-like. If their preference for IO-DO were as strong as the adults' preference for DO-IO, IO-DO would be the only object order produced in the task. We thus conclude that children are more attentive to animacy than adults.

References

Anderssen, Merete, Yulia Rodina, Roksolana Mykhaylyk, and Paula Fikkert. (2014) "The acquisition of the dative alternation in Norwegian". *Language acquisition* 21(1): 72–102.

Bates, Douglas, Martin Machler, Ben Bolker, and Steve Walker. (2015) "Fitting linear mixed-effects models using lme4" *Journal of statistical software* 67(1): 48.

Birner, Betty J. and Gregory Ward. (2009) "Information structure and syntactic structure". *Language and linguistics compass* 3(4): 1167–87.

Branigan, Holly P., Martin J. Pickering, and Mikihiro Tanaka. (2008) "Contributions of animacy to grammatical function assignment and word order during production". *Lingua* 118(2): 172–89.

Brown, Meredith, Virginia Savova, and Edward Gibson. (2012) "Syntax encodes information structure: Evidence from on-line reading comprehension". *Journal of memory and language* 66(1): 194–209. Available at: http://doi.org/10.1016/j.jml.2011.08.006.

Clifton, Charles and Lyn Frazier. (2004) "Should given information come before new? Yes and no". *Memory and cognition* 32(6): 886–95.

Cook, Vivianne J. (1975) "A note on indirect objects". *Journal of child language* 3(3): 435–37. Available at: http://doi.org/10.1017/S0305000900007315.

de Marneffe, Marie-Catherine, Scott Grimm, Inbal Arnon, Susannah Kirby, and Joan Bresnan. (2012) "A statistical model of the gramamtical choices in child production dative sentences". *Language and cognitive processes* 27(1): 86.

Eisenbeiss, Sonja. (2011) "CEGS: An elicitation tool kit for studies on case marking and its acquisition". *Essex research reports in linguistics* 60(1): 22.

Firbas, Jan. (1964) "From comparative word-order studies (thoughts on V. Mathesius' conception of the word-order system in English compared with that in Czech)". *Brno studies in English* 4: 111–26.

Gračanin-Yuksek, Martina. (2006) "Double object construction in Croatian: Arguments against Appl0". Paper presented at Formal Approaches to Slavic Linguistics 15: The Toronto Meeting, Toronto, Ontario, 12–14 May 2006.

Höhle, Barbara, Robin Hörnig, Thomas Weskott, Selene Knauf, and Agnes Krüger. (2014) "Effects of focus and definiteness on children's word order: Evidence from German five-year-olds' reproductions of double object constructions". *Journal of child language* 41(4): 780–810. Available at: http://doi.org/10.1017/S0305000913000196.

Kempen, Gerard and Karin Harbusch. (2004) "A corpus study into word order variation in German subordinate clauses: Animacy affects linearization independently of grammatical function assignment". Thomas Pechman and Christopher Habel, eds. *Multidisciplinary approaches to language pro-*

duction. Berlin: Mouton de Gruyter, 173–82. [*Trends in linguistics studies and monographs*, 157.]

Kizach, Johannes and Laura Winther Balling. (2013) "Givenness, complexity, and the Danish dative alternation". *Memory and cognition* 41(8): 1159–71. Available at: http://doi.org/10.3758/s13421-013-0336-3.

Kizach, Johannes and Thomas Mathiasen. (2013) "The dative alternation in Danish and Polish–A learners perspective". *Poznań studies in contemporary linguistics PSiCL* 49: 487–507.

Kovačević, Melita. (2004) *Croatian Kovacevic Corpus*. Available at: https://childes.talkbank.org/access/Slavic/Croatian/Kovacevic.html.

Kučerová, Ivona. (2007) *The syntax of givenness*. Ph.D. dissertation, Massachusetts Institute of Technology. Available at: http://hdl.handle.net/1721.1/39580.

————. (2012) "Grammatical marking of givenness". *Natural language semantics* 20(1): 1–30.

Kuvač Kraljević, Jelena and Gordana Hržica. (2016) "Croatian adult spoken language corpus (HrAL)". *Fluminensia: Časopis za filološka istraživanja* 28(2): 15.

Lempert, Henrietta. (1989) "Animacy constraints on preschool children's acquisition of syntax". *Child development* 60(1): 237–45.

Levin, Beth. (2008) "Dative verbs: A crosslinguistic perspective". *Lingvisticæ investigationes* 31(2): 285–312.

Malchukov, Andrej L. (2008) "Animacy and asymmetries in differential case marking". *Lingua* 118(2): 203–21.

Mykhaylyk, Roksolana, Yulia Rodina, and Merete Anderssen. (2013) "Ditransitive constructions in Russian and Ukrainian: Effect of givenness on word order". *Lingua* 137: 271–89.

Røreng, Anita. (2011) *Die deutsche Doppelobjektkonstruktion : Eine korpusbasierte Untersuchung zur relativen Abfolge nominaler Akkusativ- und Dativobjekte im geschriebenen Deutsch*. Ph.D. dissertation, Universitetet i Tromsø.

Schaeffer, Jeannette. (1999) "Articles in English child language". Paper presented at Linguistic Society of America Annual Meeting, Los Angeles, CA, 8 January 1999.

Schaeffer, Jeannette and Lisa Matthewson. (2005) "Grammar and pragmatics in the acquisition of article systems". *Natural language and linguistic theory* 23(1): 53–101. Available at: http://doi.org/10.1007/s11049-004-5540-1.

Snyder, Kieran. (2003) *The relationship between *form and *function in ditransitive constructions*. Ph.D. dissertation, University of Pennsylvania.

Stephens, Nola. (2015) "Dative constructions and givenness in the speech of four-year-olds". *Linguistics* 53(3): 405–42. Available at: http://doi.org/10.1515/ling-2015-0008.

Šimík, Radek, Marta Wierzba, and Beste Kamali. (2014) "Givenness and the position of the direct object in the Czech clause". Cassandra Chapman, Olena Kit, and Ivona Kučerová, eds. *Formal Approaches to Slavic Linguistics:*

The McMaster meeting 2013. Ann Arbor, MI: Michigan Slavic Publications, 302–18.

Titov, Elena. (2017) "The canonical order of Russian objects". *Linguistic inquiry* 48(3): 427–57.

Velnić, Marta. (2014) *The double object database*. Available at: http://linguistics-db.velnic.net:8080/user/login.

————. (2018) *Ditransitive structures in Croatian adult and child language: The role of animacy and givenness*. Ph.D. dissertation, UiT The Arctic University of Norway, Tromsø .

————. (2019) "The influence of animacy, givenness, and focus on object order in Croatian ditransitives". *Studia linguistica* 73(1): 175–201. Available at: http://doi.org/10.1111/stul.12094.

————. (forthcoming). "Acquisition of ditransitive structures in Croatian child language". Paper presented at FDSL Workshop: L1 Acquisition of syntax in the Slavic languages., Potsdam, Germany, 2–4 December 2015.

Zovko-Dinković, Irena. (2007) "Dative alternation in Croatian". *Suvremena lingvistika* 63: 65–83.

Department of Language and Literature
Dragvoll, office 5506
NTNU- Norwegian University of Science and Technology
7491 Trondheim
Norway
marta.velnic@ntnu.no / marta@velnic.net

Appendix

Table A1. Second possible order of images for 'give'.

	Given	Action	Direction
1	No given	Fox gives flower to duck	S > DO > IO[16]
2	S	**Duck** gives apple to cat	IO < DO < S
3	S and DO	**Duck** gives **flower** to horse	S > DO > IO
4	S and IO	**Fox** gives cake to **horse**	S > DO > IO
5	All	**Fox** gives **apple** to **cat**	S > DO > IO

Table A2. Second possible order of images for 'offer'.

	Given	Action	Direction
1	No given	Lion offers carrot to pig	IO < DO < S
2	S	**Pig** offers lollipop to zebra	IO < DO < S
3	S and DO	**Pig** offers **carrot** to monkey	IO < DO < S
4	S and IO	**Lion** offers sandwich to **monkey**	S > DO > IO
5	All	**Lion** offers **lollipop** to **zebra**	S > DO > IO

Table A3. Second possible order of images for 'send'.

	Given	Action	Direction
1	No given	Bunny sends parrot to turtle	S > DO > IO
2	S	**Turtle** sends puppy to elephant	S > DO > IO
3	S and DO	**Turtle** sends **parrot** to snail	S > DO > IO
4	S and IO	**Bunny** sends mouse to **snail**	IO < DO < S
5	All	**Bunny** sends **puppy** to **elephant**	IO < DO < S

[16] This image was originally supposed to have IO < DO < S order and it was illustrated that way, but during the printing process it was reversed and printed as a mirror image, resulting in the inverse orders of the characters.

Table A4. Distribution of the adult responses in the IO-animate condition.

	No-G	DO-G	IO-G	All-G
DO-IO	59% (82)	80% (56)	63% (45)	64% (43)
IO-DO	41% (58)	20% (14)	37% (26)	36% (24)
Total	348			

Table A5. Distribution of the children's responses in the IO-animate condition.

	No-G	DO-G	IO-G	All-G
DO-IO	29% (45)	24% (25)	31% (40)	26% (34)
IO-DO	71% (112)	76% (80)	69% (90)	74% (99)
Total	525			

Table A6. Distribution of the adult responses in the both-animate condition.

	No-G	DO-G	IO-G	All-G
DO-IO	97% (36)	100% (16)	100% (20)	100% (18)
IO-DO	3% (1)	0%	0%	0%
Total	91			

Table A7. Distribution of the children's responses in the both-animate condition.

	No-G	DO-G	IO-G	All-G
DO-IO	48% (11)	59% (13)	53% (10)	47% (17)
IO-DO	52% (12)	41% (9)	47% (9)	53% (19)
Total	100			

Table A8. Pairwise comparison of the givenness conditions in children when animacy is prototypical

Contrast	Estimate	SE	Z.ratio	Significance
ALL-NO	-0.298	0.311	-0.96	No
ALL-DO	0.141	0.349	0.404	No
ALL-IO	-0.556	0.325	-1.711	No
NO-DO	0.439	0.333	1.321	No
NO-IO	-0.257	0.306	-0.841	No
DO-IO	-0.697	0.348	-2.004	No

Table A9. Pairwise comparison of the givenness conditions in children when animacy is balanced

Contrast	Estimate	SE	Z.ratio	Significance
ALL-NO	0.360	0.656	0.549	No
ALL-DO	-0.427	0.686	-0.623	No
ALL-IO	0.004	0.692	0.006	No
NO-DO	-0.788	0.728	-1.082	No
NO-IO	-0.356	0.729	-0.488	No
DO-IO	0.431	0.753	0.573	No

Reviews

Maria Polinsky. *Heritage languages and their speakers*. Cambridge: Cambridge University Press, 2018. 410 p. [*Cambridge Studies in Linguistics*.] ISBN 978-1-107-04764-8. http://doi.org/10.1017/9781107252349.

Reviewed by Danko Šipka

The present book is penned by Maria Polinsky of the University of Maryland, an established syntactician, scholar of language universals, and, most importantly, student of heritage languages and their speakers. Professor Polinsky is one of the pioneers in the field of heritage languages, an emerging avenue of linguistic research. While this work reflects its author's broad interest in a wide range of languages, being a study in general linguistics it still exhibits some connections specifically with Slavic linguistics, as it contains numerous Russian and some Polish examples.

One can identify two grand linguistic narratives about heritage speakers. One is applied, and it is concerned with organizing heritage language classes and incorporating such learners into general language courses. The other is more theoretically minded, and it looks into specific linguistic properties of heritage speakers, often with an eye toward exploring broader linguistic principles. The present monograph is about the latter research direction only (as its title correctly suggests).

The field of heritage-speaker research has grown exponentially in recent years (as evidenced, among other things, by numerous papers of Polinsky's), which creates the need to summarize its achievements, reflect upon its challenges, and chart directions for further research. This is precisely what this monograph does for theoretically-oriented heritage-speaker research, being thus a welcome reaction to real-life needs.

The monograph features a logical architecture. Main concepts are defined first, the myth that languages like English cannot be heritage languages is dispelled next, followed by a review of methodology and approaches in the field. The next four chapters review research on heritage languages and their speakers in the areas of phonetics and phonology, morphology and morphosyntax, syntax, and, finally, semantics and pragmatics. At the very end, the author broadens the perspective somewhat by linking heritage speakers to endangered languages, and ends by concisely presenting the main findings of the monograph. The book is equipped with a list of references, general index, language index, preface, acknowledgements, and a list of abbreviations. I will briefly review the eight chapters of the book.

The first chapter, "Introduction" (1–37), defines the main concepts in the field. The key concept is that of heritage speaker, defined as "a simultaneous or sequential (successive) bilingual whose weaker language is the minority language of their society and whose stronger language is the dominant language of that society" (9). What is of particular importance in building a methodologically sound approach is that the author establishes the notion of baseline speakers as the ground against which one should view the performance of heritage speakers: "The notion of the 'baseline'—the language of adult first-generation immigrants that serves as the input to heritage learners—is critical for understanding what heritage speakers learn" (12). Additionally, the parameter of homeland speakers is established for non-indigenous minority languages: "For a number of languages, we can compare the *diaspora baseline*, that is the language of (first-generation) immigrants, with the language spoken somewhere in the 'old country': the *homeland language variety*" (13). Having established these main players in the process, the author discusses the main outcomes in heritage grammars (transfer from the dominant language, attrition, divergent attainment) and the main sources of divergence in heritage grammars (amount and type of input, incipient changes in the input, resource constraints, and universal principles of language structure).

The second chapter, "Heritage English" (38–75), serves to dispel the myth that the notion of heritage languages does not encompass English. Following a brief historical overview, the author presents current production data to then generalize linguistic properties of heritage English.

The next logical step in the architecture of this monograph is the discussion of methodology, which is presented in the chapter titled "How to study heritage speakers: Some observations on the methodologies and approaches" (76–113). The chapter discusses methodological considerations specific to heritage populations as well as assessment methodologies pertinent to research in this field. This chapter prepares the ground for the following four chapters (the core of this monograph), where research data gathered within the parameters of methodologies and approaches discussed here is presented.

The fourth chapter, "Phonetics and phonology" (114–63), introduces the idea of "heritage accent". It discusses phonological features of heritage speakers in production (in the heritage and the dominant language) as well as in reception. The fifth chapter, "Morphology and morphosyntax" (164–221), addresses the issues of the fate of paradigms and structural indeterminacy and ambiguity as well as morphology-encoding relationships between two constituents such as case marking and agreement. The discussion in this chapter is closely related to that in the next chapter, "Syntax" (222–90), where the issues of dependencies, binding, and word order are discussed. This core section of the present monograph concludes with the chapter titled "Semantics and pragmatics" (291–328), which contains a rather limited discussion of lexical systems and word meaning, followed by a presentation of data on prop-

ositional semantics, information structure and pragmatics, as well as social pragmatics.

The eighth chapter of this book, "Heritage speakers in unexpected places" (329–48), discusses the connection between the notion of heritage speakers and that of endangered languages. The main findings of the monograph are succinctly summarized in "Conclusions".

Heritage languages and their speakers is an important and welcome contribution to the study of heritage speakers primarily in that it gathers and systematizes various hitherto dispersed studies. The way in which this systematization is done is logical; all points are amply documented with empirical data and all claims are stated clearly.

Returning to the content that may be particularly of interest to Slavic linguists, one should mention Polinsky's analysis of the fate of the case system in heritage Russian. Discussing various changes in the case endings in heritage Russian as compared with baseline emigrant Russian, the author notes the intriguing fact that, despite the fact that heritage speakers use the endings differently than those in the baseline variety (and standard Russian), they still mark the roles in the sentence with case endings. Thus, for example, instead of (1) they will use (2), where the direct object is identical with the subject case form (i.e., it is not marked), whereas what would be the dative of recipient in baseline Russian shows up with an accusative ending *-u*.

(1) Pokazal devočke mašinku. *Baseline Russian*
showed girl_{DAT} toy.car_{ACC}
'showed the girl a toy car'

(2) Pokazal devočku mašinka. *Heritage Russian*
showed girl_{ACC} $\text{toy.car}_{DEFAULT}$

In the author's words: "Leaving the other changes in the heritage Russian case system aside, the resulting system shows one-to-one mapping between the marked case form and the semantic role recipient, even though the relevant case form is different from the form used in the baseline" (186).

Slavist readers may regret the absence from the bibliography of works by L'ubomír Ďurovič, truly indispensable for Slavic heritage-language studies: see particularly "Lingua in diaspora. Studies in the language of the second generation of Yugoslav immigrant children in Sweden" (*Slavica Lundensia* 9, 1983) and other issues of the same journal.

What is evident even from this brief review is that this field of study is dominated by those areas that are most prominent in North American linguistics at large: syntax and morphosyntax, phonetics and phonology, morphology, semantics and pragmatics (listed here in the order of prominence). What is conspicuously absent is the study of the lexicon. The author recog-

nizes this disproportion: "In comparison with other linguistic domains, heritage speakers' lexical knowledge has not received much attention, in part because it is difficult to operationalize and quantify this kind of knowledge" (292). Perhaps, along with continued research in traditional areas, a bold future direction of research on heritage speakers will be opened in the field of lexicon, where researchers would decide to do things "not because they are easy, but because they are hard". In that case, the present monograph would become extremely important not only because of what it displays (that part of its importance is evident) but also due to the lacunae to which it points.

Danko Šipka
Arizona State University
SILC, ASU, MC 0202, 851
Cady Mall, Tempe, AZ 85281, U.S.A.
Danko.Sipka@asu.edu

Ammon Cheskin. *Russian speakers in post-Soviet Latvia*: *Discursive identity strategies*. Edinburgh University Press: Edinburgh 2016. 248 pp. [*Russian Language and Society Series*.] ISBN 9780748697434 (hardcover), 9781474428507 (paper).

Reviewed by Anastassia Zabrodskaja

> How many times since I left Lebanon in 1976 to live in France have people asked me, with the best intention in the world, whether I felt 'more French' or 'more Lebanese'? And I always give the same answer: 'Both!'
>
> Amin Maalouf (2000: 3)

The above immediately came to mind when I was asked by Wayles Browne, the Review Editor, to review this book by Ammon Cheskin, a colleague I have met at several conferences. Being an Estonian Russian (according to the most often used official classification of various ethnic groups living in Estonia who happen to speak Russian as their first language) and a sociolinguist, I am often asked by laypeople and by colleagues from here and abroad whether I feel more Russian or more Estonian. The answer to this question is: neither. That was especially clear to me in April 2007 during the Bronze Nights, also known as the April Unrest and the April Events (street disorders triggered by the relocation of a Soviet war memorial, the Bronze Soldier in Tallinn). Since then I have often thought of Pavlenko and Blackledge's (2004: 18) observation that "identity becomes interesting when it is contested or in crisis", which is relevant to my research on identity construction among Estonian Russians as well as to my understanding of my multiple identity.

The last three decades have witnessed an increase in interest in the Baltic countries, in the titular languages, in Russian language use, and in identity construction among Russian minority groups. Almost all of these publications, dissertations, and research projects have "post-Soviet" in their titles. To be honest, in my work on post-Soviet Latvia I have also followed that trend. The question arises, How many more years will we researchers be talking about post-Soviet Latvia, post-Soviet Estonia etc.? If we go back in history 100 years, we see that both countries gained their independence from Bolshevist Russia in 1918. If we think about the start of the second period of independence (i.e., 1991) then we see that after 28 years it is still relevant to talk about

post-Soviet Latvia, post-Soviet Estonia, etc. But was it really relevant to talk about post-Tsarist Latvia, post-Tsarist Estonia, etc. in 1946?

My mother was born in 1946 and is now a pensioner. In 1993, when I was 12 years old, my mother then a schoolteacher, returned one evening from a school meeting organised by local authorities in the north-eastern coastal part of Estonia and said: "Nasten'ka, schools will shift to Estonian as the only language of instruction in the nearest future. Please take your studies seriously." And I did. I graduated with a gold medal from a Russian-language secondary school, I graduated cum laude from the MA program in Estonian Philology and I have taught Estonian as a Second Language for five years. But now, in March 2019, I ask myself: why do I feel more and more like a character from the film "Groundhog Day"? The only difference is that the long awaited tomorrow will never come: we still hear about post-Soviet Russians and/or post-Soviet Russian Speakers living in the Baltic countries who do not know the titular/official language, who do not study (in) the titular/official language etc. This is why I felt no enthusiasm about reviewing the book.

But to my great surprise Cheskin's book not only opens some new perspectives on "Russian Speakers in post-Soviet Latvia" but also leads me to reconsider some known facts and events in the history of the neighbouring Baltic country.

The book is clearly organized: it is composed of an Introduction (or Chapter 1), where Cheskin states his aim to highlight "that new forms of identity have been emerging in Latvia which are neither entirely 'Russian' nor entirely 'Latvian'" (2), and eight more chapters. The glossary and appendices provide detailed descriptions of the field research and facilitate an understanding of an ongoing storyline.

Chapter 2, "Discourse, Memory, and Identity", presents a clear contextual background, where it is explained, based on a number of approaches, why the "Other" (who is naturally Russian) is dangerous for Latvians. The same topic continues in Chapter 3, "Latvian State and Nation-Building", where again and again the reader is reminded of this threat because "in the case of Latvia, history and memory were clearly utilised as discursive tools that could help to construct a 'core' group of 'Latvians'" to be "further unified through contrast to the external 'other': 'the Russians'" (63). This chapter reveals the hypocrisy of the Popular Front of Latvia in late Soviet times and the future official actions of policy-makers, which tied the national language to history and established it as an expression of pride and resistance to Soviet rule, in a clear manner. To understand, it is enough to read just one sentence of the discussions of the citizenship law in the PFL's newspaper *Atmoda* in June 1990: "Nationalism is not intended as discrimination against other people, but rather as a cultural principle, an external boundary to protect you from others, and an internal boundary to protect others from you" (47).

Chapter 4, "Russian-Language Media and Identity Formation", investigates discourses of a leading Russian-language daily, *Chas* 'the Hour'. It is not surprising that, according to the concept of the logic of equivalence (Laclau and Mouffe 1985) mentioned by Cheskin (32), Russian-language media discourses mirror the mainstream ones, and not always as distorted mirrors. This chapter is also instructive to read as it provides examples of who "nashi" (ours) are and how Russian-speaking elites justify their support of official discourses of linguistic domination over local Russian-speaking minorities.

In Chapter 5, "Examining Russian-Speaking Identity from Below", the author elaborates on grass-roots ideological discourses that address political topics and aspects of the status of the Russian language, Russians, Russian-speakers etc. It is no great surprise that older respondents are not happy to be called either "Russian-speaking" or "non-citizen" (see page 106), which is not the case for the younger generation, who actually talk about how strange Russia is (108). This situation is logical because, in the case of the older generations of Russian-speakers or Russians living in the Baltic countries, the situation is seen widely as illegitimate. Their understanding is that they came to this territory when it was one country, the Soviet Union, and it still remains unclear to this group of people how it is possible to become occupiers overnight. It is clear, based on the example of Latvia and Estonia, that these grass-roots discussions only reinforce the top-down approach, which assigns greater importance to Latvian (or Estonian), protecting it tirelessly against the forces of the everlasting enemy: the "Other". But for younger generations who were born in Latvia a shift in perception has occurred. As Cheskin notes: "Naturally it is important to bear in mind the selection bias of this research, which focused on educated people who were linguistically advantaged" (128), and I would add, "and from Riga" (the capital of Latvia). This chapter should be taken as just one case study showing how only certain Russians undergo the processes of linguistic, ethnic, and national identity formation.

Chapter 6, "The 'Democratisation of History' and Generational Change", makes an effort "to move away from a strict reliance on elite discourses and the analysis of such discourses" (130). In this chapter, the figures are poorly presented in: the colors used are almost the same, and it is hard to distinguish grey from the light grey and the dark grey. But if you ignore those technical obstacles, then, as in Blok's poem, "And an eternal fight, we only dream of rest" this "eternal fight" regarding "occupation—annexation" between Latvian and Russian Federation official discourses is nicely illustrated via bottom-up positions different age groups hold in their historical interpretations.

When I read extracts from "semi-structured interviews … conducted with six members of the Latvian Parliament" (see more on page 152) in Chapter 7, "The Primacy of Politics? Political Discourse and Identity Formation", I felt like I was in the 1979 Soviet science fiction art film "Stalker", in which getting out of the "Zone" is only possible if there is cooperation against the "Other"

and constant protection of Latvians. Déjà vu: these repeated discourses painfully remind me of discourses from the early 1990s. As for Chapter 7's content, I would criticise the title of the section "Rising political tensions 2010–14", as there is not much mention of 2014, an extremely important turning point in relations between the Baltic countries, the Russian Federation, and the West: EU, NATO, and the U.S.

Chapter 8, "The Russian Federation and Russian-Speaking Identity in Latvia", clearly needs to be revised in a possible second edition, adding more contemporary data on political discourse. The political changes in Latvia and in the "compatriot policy" of the Russian Federation that have occurred have already had and will have more far-reaching effects than just changes in compatriot consolidation and identity and/or political discourse construction. In its current form, Chapter 8 does not reflect the possible dynamics in these processes.

Cheskin tells the reader: "It is argued, through the course of this book, that it is very difficult for Russian speakers to find a legitimate place within Latvian discourses because the process of othering is central to the formation of Latvian identity in the first place" (33). In the concluding Chapter 9, which is optimistically entitled "A Bright Future?" on the last page of the book Cheskin states: "In light of the hardening discursive positions of the respective Latvian and Russian states, Latvia's Russian speakers will continue to be faced with contradictory identity pressures." If this conclusion is weak, it is not the author's fault. He has attempted a Sisyphean task, exploring "discursive identity strategies" of "Russian speakers in post-Soviet Latvia". His study suggests that the long awaited tomorrow will never arrive. An explanation of the situation not only in "post-Soviet" Latvia but also in "post-Soviet" Estonia (as well as in many other "post-Soviet" countries, which are outside the scope of this review) might be provided by Robert Rozhdestvensky's (1992) poem:

Dlja čeloveka nacional'nost–
I ne zasluga,
I ne vina.
Esli v strane
Utverždajut inače,
Značit,
Nesčastna èta strana!

For a person, nationality
is neither merit
nor fault.
If in a country
it is argued differently
this means this country is unhappy.

References

Laclau, Ernesto and Chantal Mouffe. (1985) *Hegemony and socialist strategy: Towards a radical democratic politics*. London: Verso.

Maalouf, Amin. (2000) *On identity*. Trans. by Barbara Bray. London: Harvill Press.

Pavlenko, Aneta and Adrian Blackledge. (2004) *Negotiation of identities in multilingual settings*. Clevedon, U.K.: Multilingual Matters.

First affiliation:
Baltic Film, Media, Arts and Communication School
Tallinn University, Narva mnt 27, room N430
Tallinn 10120, Estonia

Second affiliation:
Institute of Estonian and General Linguistics
University of Tartu, Jakobi 2
Tartu 50090, Estonia
anastassia.zabrodskaja@gmail.com