

CONTEXT-INDUCED GRAMMATICALIZATION OF THE INDEFINITE ARTICLE *IN STATU NASCENDI*^a

Branimir Stanković*

University of Niš, Faculty of Philosophy, Niš, Serbia

ORCID iD: Branimir Stanković

<https://orcid.org/0000-0002-5542-544X>

Abstract

Based on the mandatory use of indefinite-determiner-like elements in Serbian, a language with no categories of definite and indefinite article, the paper proposes the hypothesis that the language universal grammaticalization of the cardinal number *one* into an indefinite article is triggered by specific types of contexts, rather than simply by the speakers' intention to disambiguate between the definite or indefinite reading of the 'bare' noun phrase. Given the results of a pragmatic felicitousness test of 35 Serbian native speakers' intuition, we demonstrate that the exhaustiveness of the nominal description directly influences the obligatoriness of the use of the 'determiner'—the less informative the description is, the more compulsory the use of the indefinite determiner becomes. We argue that this actually represents the *spiritus movens* in the grammaticalization of indefinite articles in the languages with this category, via the subsequent process of gradual context-induced reinterpretation.

Key words: grammaticalization, indefinite article, Serbian language, mandatory determiners, context-induced reinterpretation.

ГРАМАТИКАЛИЗАЦИЈА НЕОДРЕЂЕНОГ ЧЛАНА ИНДУКОВАНА КОНТЕКСТОМ *IN STATU NASCENDI*

Апстракт

На основу обавезне употребе језичких израза који наликују неодређеним детерминаторима у српском, језику који нема категорије одређеног и неодређеног члана, рад предлаже хипотезу да језички универзална грамаатикализација кардиналног броја *један* у неодређени члан бива покренута специфичним типовима контекста, пре него пуком намером говорника да раздвозначе између одређеног и неодређеног читања „голе” именичке синтагме. На основу резултата спроведеног теста прагматичке подесности на интуицији 35 изворних говорника

^a The article was presented at the *Language, Literature, Process 2023 Conference* at the Faculty of Philosophy, University of Niš, Serbia.

* Corresponding author: Branimir Stanković, University of Niš, Faculty of Philosophy, Ćirila i Metodija 2, 18000 Niš, Serbia, branimir.stankovic@filfak.ni.ac.rs

српског, показујемо да исцрпност номиналног описа директно утиче на обавезност употребе „детерминатора” – што је опис мање информативан, то је већа обавезност употребе неодређеног детерминатора. Заговарамо идеју да ово, у ствари, представља покретачку силу за граматикализацију неодређеног члана у језицима са овом категоријом, посредством каснијег процеса постепене реинтерпретације индуковане контекстом.

Кључне речи: граматикализација, неодређени члан, српски језик, обавезни детерминатори, реинтерпретација индукована контекстом.

INTRODUCTION

Although Serbian, like most languages, displays a regular set of indefinite pronoun items (derived from the interrogative/relative pronouns by prefixing with *ne-*), when introducing new, epistemically proximal animate¹ nominal referents with an **impoverished** description to the discourse out-of-the-blue, its speakers are forced to use the number *jedan*, meaning ‘one,’ with its non-cardinal, indefinite interpretation, or else the nominal phrase would receive unambiguous definite interpretation:

- (1) Ukus umami je otkrio {#(jedan)/(?neki)} japanski naučnik.
 Njegovo ime je...
 flavor umami AUX discovered one some Japanese scientist
 his name is
 “The umami flavor was discovered by {a / (?some)} Japanese scientist. His name is...”

In (1), the noun phrase (NP) *japanski naučnik* (*Eng.* Japanese scientist) is not part of interlocutors’ discourse-model and its referent is epistemically proximal to the speaker, indicated by the fact that the locutor is about to reveal his name in the following sentence. Even though Serbian (as the case is with most Slavic languages except for the intensively ‘balkanized’ Macedonian and Bulgarian), does not have the categories of indefinite and definite article, its speakers are obligated to use the indefinite-determiner-like element *jedan*/‘one’, or else the NP would not be pragmatically felicitous (marked with the # symbol and the appropriate bracketing). On the other side, the regular indefinite pronoun *neki*/‘some’ in this context is somewhat inappropriate, as it would implicate epistemic distance from the locutor, which is contrary to the remainder of the discourse, where more information is to be supplied about the nominal referent. As already noted, the only available interpretation of the NP without

¹ We thank the anonymous reviewer for the remark that the observed phenomenon is limited to animate referents (*cf.* *Petar je kupio #jedna kola.* ‘Peter bought a car’). This asymmetry suggests that animacy also plays an important role in the analysed processes, an issue that should certainly be addressed in future research.

the ‘determiner’ would be definite, i.e., that its referent is familiar to the co-locutor.

Now let’s observe what happens in case the NP does not have an impoverished description—as we have labelled it—meaning, consisting only of a relational adjective² modifying a noun,³ but an enriched one, such as *japanski naučnik sa Tokijskog kraljevskog univerziteta* ‘(a) Japanese scientist from the Tokyo Imperial University’:

- (2) Ukus umami je otkrio japanski naučnik sa
Tokijskog kraljevskog univerziteta.
flavour umami AUX discovered Japanese scientist from
Tokyo Imperial University
‘The umami flavour was discovered by a Japanese scientist from Tokyo
Imperial University.’

Immediately the NP in such context becomes pragmatically well-formed, although lacking any indefinite-determiner-like element, unlike the situation in (1). Of course, this does not mean that the speaker could not utilize *jedan* ‘one’ for introducing the discourse-new NP. But, contrary to (1), in (2) he/she is not obliged to do so, which is to be expected from an article-less language.

The presented asymmetry leads us to the working hypothesis to be verified in the paper, that the less informative the animate nominal description is in Serbian, the more compulsory the use of the indefinite determiner becomes. We will argue that this actually represents the *spiritus movens* in the grammaticalization of indefinite articles in the languages with this category, via the subsequent process of gradual context-induced reinterpretation (Heine, Claudi & Hünnemeyer, 1991), a concept to be explained in the following chapter.

The structure of the paper is as follows. First, we will briefly present Heine’s (1997) five-stages grammaticalization model and the mentioned context-induced reinterpretation approach, as well as the relevant previous research on article-like use of the cardinal number *one* in Slavic languages. Afterwards, we will elucidate our hypothesis and demonstrate

² We are employing a relational adjective intentionally, because this adjective class cannot mark the distinction between the Serbian definite and indefinite adjectival aspect. In case a descriptive adjective was used, the indefinite adjective form could unambiguously trigger indefinite reading, a scenario we are trying to avoid.

³ It goes without saying that the use of an indefinite marker is necessary for obtaining an indefinite reading of a ‘bare’ noun:

i. Ukus umami je otkrio {#(jedan)/(?neki)} naučnik. Njegovo ime je
Kikunae Ikeda.
flavor umami AUX discovered one some scientist his name is
Kikunae Ikeda

‘The umami flavor was discovered by {a / (?some)} scientist. His name is Kikunae Ikeda.’

the methodology of our survey, followed by a chapter dedicated to discussion of the results and our final remarks.

PREVIOUS RESEARCH

There is an extensive body of research on the issue of grammaticalization of the cardinal number *one* into an indefinite article both in languages with this category, as well as in article-less ones, which happens to be the vast majority of Slavic languages. Limited with space, we will present only a portion of the relevant literature, but will direct the readers to further exploration of the surveys referenced in the presented studies. It is important to notice that the crucial difference between the approach taken in our paper and in previous literature is the **obligatoriness** in speech. All earlier studies concentrate on the problem **can** the cardinal number *one* be utilized as a grammatical marker in certain context(s), whereas we focus on the question – in which contexts **must** speakers use *one* to get the appropriate indefinite reading or else the NP receives an unambiguous definite interpretation?⁴ This distinction is not a trivial one, as it brings serious theoretical consequences, leading to the assumption that what we are witnessing in Serbian *in statu nascendi* is actually a grammaticalization path taken in all languages with the category of indefinite article (see *Hypothesis and methodology* and *Discussion* sections for an elaboration on this).

Theoretical Background

Moravcsik (1969) estimates that the numeral item ‘one’ is an optional indefiniteness marker crosslinguistically, assuming that certain traits of the indefinite article can be reduced to and interpreted as properties of the cardinal it originates from. She also points out the prosodic changes that it usually undergoes, resulting in cliticization of the novel indefinite determiner in certain languages (*cf.* Belaj & Matovac’s (2015) account on this situation in Croatian below).

In his monograph “Cognitive Foundations of Grammar,” Heine (1997) proposes the following five-stage model, which can be employed both as a synchronic implicational scale (in the sense that an indefinite article for each stage also has (or may have) the properties of all preceding stages, but never vice versa) and as a diachronic evolution apparatus aiming at describing the grammaticalization processes involving the item *one* and its cross-linguistic counterparts (Heine, 1997, pp. 72–75):

⁴ For instance, Heine (1997, p. 69) points out that even in cases when a language has a grammaticalized indefinite article, it is not necessarily used in all NPs with indefinite reference.

(3) numeral > presentative marker > specific marker > nonspecific marker > generalized article

At **stage 1**, the item *one* can only function as a **numeral**. This is the situation, for instance, in Swahili, in which the cardinal number *moja* ‘one’ is used exclusively for this purpose, and never as a presentative or a specific marker (see the explanation in the following two paragraphs). In the latter case, the indefinite specific-denoting NP is left ‘bare,’ without this or any other kind of marker.

At **stage 2**, the item *one* functions as a **presentative marker**, which introduces new referents presumed to be unknown to the collocutor and taken up as definite and relevant in subsequent discourse, as witnessed, for instance, in the Gurage language Soddo-Goggot of Ethiopia (Hetzron, 1977, p. 56, as cited in Heine (1997)). At an early stage of grammaticalization, the use of *one* as a presentative marker article could be limited strictly to the beginning of a narrative discourse. A perfect example is the Uto-Aztecan Western Tarahumara language, in which the NP referents are accompanied with the presentative article only at the beginning of a tale, but all other specific NPs remain unmarked⁵.

The next, **stage number 3** pertains to situations in which the cardinal number *one* is regularly used for all (or vast majority of) singular indefinite **specific** nominal referents of countable nouns, irrespective of their role in subsequent discourse (the label *specific* could be defined as discourse-participants familiar to the speaker but presumed to be unfamiliar to the hearer). Givón (1981, p. 36) demonstrates that in Street Hebrew, the item *exad* ‘one, a’ is used in indefinite specific contexts, but not in situations in which solely membership to a certain type wants to be highlighted:

- (4) ba hena (#exad) ish etmol, lo (#exad) isha!
 came here a man yesterday not a woman
 “A man came here yesterday, not a woman!”

At **stage 4**, indefinite articles function as **nonspecific markers**, determining NPs “whose referential identity neither the hearer nor the speaker knows or cares to know” (Heine, 1997, p. 73). This level of grammaticalization is present in modern English, German, Dutch, most Romance languages, Punjabi and Chinese, to name a few.

Finally, at **stage 5**, the **generalized article** occurs both with singular and plural nouns, and with count and mass nouns, as well. Exceptions from this could be situations in which the nominal referent is already definite, then with proper nouns, or with predicate nouns “defining members

⁵ Đurić (2023) presents an elaborate contrastive diachronic corpus analysis of *one* used as a presentative marker at the beginning of Serbian and (Serbian translations of) Russian folk tales.

of an ethnic, professional, or some other class ('He is Swiss')" (*ibid.*). This final stage of indefinite article grammaticalization is present, for instance, in modern Spanish and Portuguese.

The author underlines that the presented stages should not be comprehended as discrete entities, and that the development from stage 1 to stage 5 is continuous, involving different kinds of overlaps. As had been proposed in his Overlap Model (Heine, 1993), grammaticalization always involves intermediate stages, in which the linguistic expression "can be interpreted alternatively with reference to the earlier and the later structure" (Heine, 1997, p. 74).

In a similar manner, Heine et al. (1991) argue that grammaticalization based on metaphorical and/or metonymical transfer from one cognitive domain to another only appears to be realized in discreet steps, but is actually gradual and continuous in its nature. This is illustrated with the full development of the Ewe lexeme *meɣbé*/'back'—which can denote a body part (OBJECT), a locative content used as an adverb or a postposition (SPACE), a temporal content (TIME), and finally, a mental, cognitive trait (QUALITY)—by presenting various examples in which this lexical item can refer simultaneously to more than one of the categories OBJECT, SPACE, TIME and QUALITY. To account for this and similar phenomena, the authors introduce the notion of "context-induced reinterpretation," which consists of the following three stages:

Stage I: In addition to its focal or core sense *A*, a given linguistic form *F* acquires an additional sense *B* when occurring in a specific context *C*. This can result in semantic ambiguity since either of the senses *A* or *B* may be implied in context *C*. Which of the two senses is implied usually is, but need not be, dependent on the relevant communication situation. It is equally possible that the speaker means *A* and the hearer interprets him or her as implying *B* or that the hearer understands *B* whereas the speaker intends to convey *A*.

Stage II: The existence of sense *B* now makes it possible for the relevant form to be used in new contexts that are compatible with *B* but rule out sense *A*.

Stage III: *B* is conventionalized; it may be said to form a secondary focus characterized by properties containing elements not present in *A* [...] with the effect that *F* now has two 'polysemes,' *A* and *B*, which may develop eventually into 'homophones.'

(Heine et al., 1991, pp. 71–72)

As will be presented in the *Hypothesis and methodology* section of the paper, we will argue that the grammaticalization of the numeral into a presentative marker and subsequently into a specific marker is triggered once a given linguistic form *F₁*, a bare noun or a nominal phrase with an impoverished description, can only receive a definite interpretation in a specific context *C*—in our case, when the nominal referent is familiar to

the speaker, but unfamiliar to the hearer. As a consequence, the locutor is forced to employ some sort of an indefinite determiner-like element, resulting in the mentioned *Stage I*, in which a given linguistic form F_2 (the item *jedan*/'one'), in addition to its focal or core sense *A* (numeral), acquires a sense *B* (presentative marker/specific marker) when occurring in the context *C*.

Grammaticalization of the Indefinite Article in Slavic languages

In their study “The Changing Languages of Europe”, Heine & Kuteva (2006) provide a thorough examination of the definiteness and indefiniteness grammaticalization status of the nominal phrase in all European languages, including Slavic ones. They report that, unlike Ukrainian and Belarussian, Russian⁶ and Colloquial Polish have already reached the presentative marker stage, that Czech, Colloquial Serbian and Colloquial Croatian are somewhere between stages 2 and 3 (presentative marker and specific marker), while Bulgarian, Eastern Macedonian and Upper Sorbian have completely reached stage 3. It is important to notice that the authors consider language contact as a major contribution for the given data:

[T]hose eastern European languages which are geographically close to western European languages (e.g. Sorbian) as well as to both west European languages and the south European language Greek (i.e. Albanian, Bulgarian, Macedonian) are also languages with a clearly developing indefinite article, which may well represent a stage 2 or a stage 3 situation.

(Heine & Kuteva, 2006, p. 131)

Hwaszcz and Kędzierska (2018) also present an updated informative cross-linguistic state of the art overview of grammaticalization of the indefinite and definite article in all Slavic languages (we kindly direct the readers to the papers presented therein), but with special focus on the situation in Polish. Based on 53 native speakers' judgments, and the corpus analysis of 20,000 sentences containing the item *jeden*/'one', the authors conclude that this numeral has already reached the specific marker stage in Polish, “with some occasional uses leaning towards the non-specific marker stage” (Hwaszcz & Kędzierska, 2018, p. 93). Following Heine & Kuteva (2006), Hwaszcz and Kędzierska also explain the attested grammaticalization in Polish as enhanced by language contact with languages such as English and German, which both have the categories of indefinite and definite article.

⁶ The authors emphasize that in Russian, this is limited to headlines of scientific articles and—sometimes—to presentative contexts.

A paper to address almost the same topic we are dealing with is “On the article-like use of the indefinite determiners *jedan* and *neki* in Croatian and other Slavic languages” by Belaj and Matovac (2015). The theoretical approach that the authors employed is functional–typological, with minor elements of the Cognitive Grammar framework. The crucial criteria for contrasting the level of grammaticalization of the indefinite determiners in different Slavic languages is Heine’s (1997) five-stage model, presented in detail in the previous section of this paper. Firstly, Belaj and Matovac examine the article-like usages of *jedan*/‘one’ in Bulgarian, Macedonian and Upper Sorbian, given the fact that those three languages have reached the highest degree of grammaticalization. In the following section, they turn to the situation in Croatian. Based on their native speakers’ intuition, the authors demonstrate that in this language *jedan*/‘one’ can be used as a numeral, an adjective,⁷ or an indefinite determiner with article-like use. In addition, they state that this item “can be used to determine any nominal within a sentence, even ones used in predicative constructions”⁸ (Belaj & Matovac, 2015, p. 9). When it comes to the distinction between *jedan* functioning as a numeral or an adjective and as an indefinite determiner, the authors claim that in the latter case, the analysed element is not stressed, therefore “[...] it behaves as a proclitic, forming a single accentual unit with the following word” (*ibid.*). They find this to be a common trait with the indefinite article-like use of the pronominal item *neki*/‘some’. Curiously enough, the authors recognize the fact that *jedan* and *neki* “used as indefinite articles will be stressed when followed by enclitics that are always unstressed” (*ibid.*, p. 10).

Leaving the unlikely possibility that there might be fundamental prosodic differences between the situations in Croatian and Serbian,⁹ it

⁷ This is illustrated by the following example, which more seems to be a pronominal, than an adjective use (the English translation is quoted as provided in the paper):

(i) Jedni majstori su tek otišli, a drugi već dolaze.

“No sooner had one workmen gone than the others arrived.” (Belaj & Matovac, 2015, p. 9)

⁸ This statement is exemplified with the next sentence, which looks more like a non-referential than a genuine referential use:

(ii) Ti si jedna velika budala.

“You are a (lit. one) big fool.” (*ibidem*)

⁹ As pointed out by the authors themselves, Croatian, Serbian, Bosnian and Montenegrin are based on the same, Štokavian dialect, so they do not differ significantly in the indefinite determiner use of *jedan* and *neki*, followed by the remark that it is beyond the scope of their paper to identify the differences between them. Interestingly, they cite Friedman’s (2000, p. 196) claim that *jedan* is more grammaticalized in Serbian than in Croatian, which is explained as a result of the more intensive language contact Serbian has with the languages of the Balkan Sprachbund, especially Bulgarian. If this is correct (and we believe it is), one would expect more progressed cliticization to be found in Serbian, which is contrary to the assumption made above that the potential prosodic differences are such

seems that in the presented paper, sentence intonation based on information structure has been mistaken for word stress. Namely, the analysed elements in the numeral and in the so-called adjective use allow topicalization or focalization, and can bear a contrastive intonation, as well. All of these might be intonationally released separately from the rest of the NP. On the other side, when functioning as an indefinite determiner, *jedan* and *neki*¹⁰ are always part of the same, topical or focal, partition together with the remainder of the nominal phrase, or in some examples, they are simply backgrounded. We assume that these latter cases give the misleading impression that the analysed elements are proclitics¹¹.

In their concluding remarks, Belaj and Matovac (2015) point out that the analysed grammaticalization process still does not offer enough evidence for the claim that modern Slavic languages display the category of indefinite articles. Nevertheless, they state that their data proves that the situation presented by Heine and Kuteva (2006) needs to be reconsidered, “particularly the claims stating that the Croatian ‘one’ still has not fully reached the specific indefinite marker stage of grammaticalization” (*ibid.*, p. 16).

We finish with the brief literature overview and the description of the theoretical background of our research. In the following chapter we turn to our hypothesis and the methodology applied for its verification.

HYPOTHESIS AND METHODOLOGY

At the beginning of the chapter dedicated to previous research, we have already touched upon the fundamental difference that separates our approach to grammaticalization of indefiniteness from the one(s) present in linguistic literature. To our knowledge, all papers investigating this process are mostly concerned with the question **can** the numeral *one* be utilized in a certain language as a presentative marker, a specific marker, a non-specific marker or a generalized article, and applying the diagnostics to determine which stage of the grammaticalization path proposed by

that cliticization is only present in Croatian. In addition, we would expect for this process to be already finished in Bulgarian and Macedonian, which certainly is not the case.

¹⁰ The authors suggest that the distinction between the use of *jedan* and *neki* is based on specificity, stating that *neki* refers to non-specific referents only, while *jedan* can refer both to specific and non-specific referents. Once more, we leave the possibility that there might be serious differences between Croatian and Serbian, because in Serbian, *neki* can be used for specific referents, as well, but only when they are epistemically distant from the locutor, as shown in (1) and (2).

¹¹ This also leads us to the conclusion that there are actually no unsystematic differences between the examples in which *jedan* and *neki* precede enclitics and stressed elements, as suggested by the authors.

Heine (1997) has been reached in a given language. In our paper, we attempt to shed light on the following issue: in which contexts **must** the speaker use an indefinite determiner-like element in order to get the appropriate indefinite reading, or else the nominal phrase receives an unambiguous definite interpretation?

The difference in the approach is not trivial by any chance. Both indefinite and definite article are very often observed as superfluous, redundant categories. Some half a century ago Beckmann (1972) assessed that the article was a redundant morpheme—there are more languages in the world lacking articles, but yet there is no difficulty in determining the nominal referents when communicating in Chinese, Russian or Serbian¹². Moreover, the article is “discarded as unnecessary and wasteful in newspaper headlines and telegrams,” even in the cases when “the grammar of the language would require its use” (Beckmann, 1972, pp. 165–166). Heine (1997) also demonstrates that the indefinite article is not necessarily utilized in all NPs even in situations when this category has been fully grammaticalized in a given language. Considering all this, one must question herself/himself: how come articles even emerge in world languages and why would they, given their superfluous nature?

In addition, it is interesting to reassess the language contact explanation proposed in previous papers (Heine & Kuteva, 2006; Belaj & Matovac, 2015; Hwaszcz & Kędzierska, 2018, a. o.). That language contact (substrates, adstrates, and/or superstrates) can enhance grammatical change does not seem controversial at all, but this approach is incapable of answering to the questions: What could have triggered the grammaticalization of the indefinite article in Greek, English or German in the first place? What has happened in isolated languages in which this process has propelled spontaneously?

In order to offer an explanation for these issues, we hypothesize that certain types of contexts are responsible for initiating the grammaticalization process. Consider the following example, a typical beginning of a narrative discourse, in this case, of a tale, which is often estimated to be the very first step to the grammaticalization of *one* into a presentative marker:

- (5) Nekada davno, živela #(jedna) kraljica.
 at one time long ago lived one queen
 “Once upon a time, there lived a queen.”

In case the narrator introduces a ‘bear’ noun to the discourse (*kraljica* – *Eng.* queen), the nominal phrase is pragmatically infelicitous, unlike the situation when accompanied with the indefinite determiner-like

¹² We leave aside the fact that cross-linguistically there are various modes for marking the referential status of the NP other than the definite and the indefinite article: word order, information structure, case marking etc.

element *jedna*/'one'. Nevertheless, the presence of the same item is not obligatory when the nominal description is comprehensive, thus providing the hearer with more elaborate instructions for identifying the referent:

- (6) Nekada davno, živila (jedna) lepa, ali ohola kraljicakoj je
 at one time long ago lived one beautiful but haughty queen who AUX
 mučila svoje podanike.
 tortured her vassals
 "Once upon a time, there lived a beautiful, but haughty queen that tortured
 her vassals."

The hypothesis that emerges out of these examples is pretty straightforward: the less informative the animate nominal description is, the more compulsory the use of the indefinite determiner becomes, and vice versa, the more elaborate/informative the description is, the less mandatory its use becomes.

Given the fact that in (5) and (6) we are dealing with a beginning of a folk tale, the discourse model shared by the interlocutors is empty, implicating that the introduced NPs should be interpreted as indefinite. Nevertheless, if we take into consideration all examples other than typical beginnings of a narrative discourse, the referential status of elaborate NPs modified by prenominal adjectives and postnominal relative clauses and/or preposition phrases is not unambiguously indefinite. On the contrary, they are usually ambiguous between indefinite and definite interpretation, whereby the latter reading is possible for discourse-old referents that are cognitively distant, located at the very end of Gundel, Hedberg & Zacharski's (1993) Givenness Hierarchy, labelled by the authors as *type identifiable*:

- (7) **in focus** (*it*) > **activated** (*that; this; this N*) > **familiar** (*that N*) > **uniquely identifiable** (*the N*) > **referential** (indefinite *this N*) > **type identifiable** (*a N*) (Gundel et al., 1993, p. 275)

The authors assume that there are six cognitive statuses relevant to the form of referring expressions in natural language discourse, from nominal expressions *in focus*, which are easily cognitively accessible and usually marked with a personal pronoun (*it*), to the *type identifiable*, which are least accessible, meaning, that the addressee is only able to access a representation of the type of object described by the expression, and consequently they are marked with an indefinite determiner (*a N*). What we are suggesting is that in article-less languages not all nominal expressions of the latter type (*type identifiable*) are subject to the same level of mandatory indefiniteness marking during the process of grammaticalization – it is initiated with 'bare' nouns, and then it gradually extends to nouns modified by a prenominal adjective, and/or postnominal preposition phrase / relative clause.

As already explained, we argue that the grammaticalization process is induced once a given linguistic form F_1 , a nominal phrase with an impoverished description, can only receive a definite interpretation in a specific context C , i.e. when the nominal referent is known to the speaker, but not to the hearer. Consequently, the speaker is forced to use some indefinite determiner-like item, resulting in Heine et al.'s (1991) *Stage I* of context-induced reinterpretation, in which a given linguistic form F_2 (the item *jedan*/'one'), in addition to its focal or core sense A (numeral), acquires a sense B (presentative/specific marker) when occurring in the context C . Once initiated in contexts containing 'bare' nouns, the non-cardinal use gets reinterpreted progressively in all other, more elaborate nominal descriptions.

In order to verify the postulated hypothesis, we performed an online pragmatic felicitousness test on 35 native speakers of the Serbian language. There were 12 male participants (34.28%) and 23 female participants (65.71%). The age of the participants ranged from 19 to 53 years, with an average age of 27.24 years. The majority of them, more precisely 27 (77.14%), had linguistic education or were studying a philological discipline at the time of conducting the experiment, while the rest (22.86%) were individuals without any formal education in linguistics. 16 individuals (45.71%) were native speakers of the Prizren-Timok dialect, 11 subjects (31.43%) were from the Kosovo-Resava dialect area, 4 individuals (11.43%) were speakers of the Smederevo-Vršac dialect, and the same number of participants came from the Šumadija-Vojvodina dialect area.

The subjects were supposed to rate how well-formed do the target sentences appear to them on a Likert scale ranging from 1 to 5, imagining that the referents of the critical NPs, underlined in the text, had not been mentioned before. The stimuli were divided in 3 groups consisting of 6 target examples: (1) 'bare' nouns; (2) nouns modified by a prenominal adjective; and (3) nouns modified by a prenominal adjective and a postnominal preposition phrase or a relative clause. In 7(a-c), the corresponding examples are presented for each of the three levels of the independent variable:

- (7.a) Naučnik je otkrio ukus umami.
“(A) Scientist discovered the umami taste.”
- (7.b) Japanski naučnik je otkrio ukus umami.
“(A) Japanese scientist discovered the umami taste.”
- (7.c) Japanski naučnik sa Tokijskog kraljevskog univerziteta je otkrio ukus umami.
“(A) Japanese scientist from the Tokyo Royal University discovered the umami taste.”

The critical sentences were randomized and accompanied with 24 distractors. In the following chapter, we bring the discussion of the results and our concluding remarks.

RESULTS DISCUSSION AND CONCLUSION

After obtaining the results of the performed research, the raw data immediately suggested that our initial intuition was correct. The arithmetical means for the estimation grades were 1.37, 2.94, and 4.77, for the 3 levels of our independent variable ('bare' noun; adjective + noun; adjective + noun + preposition phrase / relative clause), respectively. These results unequivocally demonstrated a clear increase in the acceptability of nominal phrases in indefinite contexts (but without the employment of indefinite pronouns or the indefinite determiner use of the cardinal number *jedan*/'one') when the corresponding description is enriched with pronominal modifiers. Moreover, these nominal phrases become fully acceptable in the presence of both pronominal and postnominal modification/relative clauses. We then conducted Spearman's correlation test, and received an unambiguous confirmation of the hypothesis, with the fairly high coefficient $r_s = 0.94631$. All things being equal, the more elaborate the animate-denoting nominal description, the more acceptable it is when introduced into the discourse out-of-the-blue unaccompanied by any type of indefinite-like determiner.

Our experiment design involved evaluating the pragmatical well-formedness of sentences with various NPs lacking indefinite determiners. This means that our subjects were **not** exposed to examples **with** the item *jedan*, so future research should try to establish what is the ratio between the acceptability of the two groups of sentences (with and without the indefinite determiner). One could assume that the difference between the 3 stimuli pairs ('bare' noun vs. *jedan* + 'bare' noun; adjective + noun vs. *jedan* + adjective + noun; adjective + noun + preposition phrase / relative clause vs. *jedan* + adjective + noun + preposition phrase / relative clause) is going to monotonically increase.

Finally, the novel insights demonstrated in this paper provoke for at least one more potential working hypothesis to emerge, and that is the assumption that the grammaticalization of the cardinal number *one* into an indefinite article is possibly universal, crosslinguistically triggered by specific types of contexts (low informative nominal descriptions), rather than simply by the speakers' intention to disambiguate between the definite/indefinite reading of the NP. There are hints that similar processes could be identified also within the area of grammaticalization of definiteness (s. Stanković 2023), leading to the conclusion that our, somewhat improved, context-induced reinterpretation analysis might be a powerful grammaticalization explanatory model.

Acknowledgement. The research was supported by the Science Fund of the Republic of Serbia, GRANT No 1589, Project title ClaCoLaSI.

REFERENCES

- Beckmann, P. (1972). *The structure of language: A new approach*. Boulder: Golem Press.
- Belaj, B. & Matovac, D. (2015). On the article-like use of the indefinite determiners *jedan* and *neki* in Croatian and other Slavic languages. *Suvremena lingvistika*, 79, 1–20.
- Đurić, Lj. (2023). Leksema 'jedan' u egzistencijalnim rečenicama na početku teksta: analiza na korpusu narodnih bajki [The lexeme *jedan* 'one' in existential sentences at the beginning of the text: folk tales corpus analysis]. Presented at the *Language, Literature, Process* conference, Faculty of Philosophy in Niš, April 2023.
- Givón, T. (1981). On the development of the numeral 'one' as an indefinite marker. *Folia Linguistica Historica*, 2/1, 35–53.
- Gundel, J.K., Hedberg, H. & Zacharski, R. (1993). Cognitive Status and the Form of Referring Expressions in Discourse. *Language*, 69(2), 274–307.
- Friedman, V. (2000). Observations on the use of *jedan* as a marker of indefiniteness in Serbian/Croatian". In: Gorup, R. & Rakić, B. (Eds.): *In a Foreign Harbor: Essays in Honor of Vasa D. Mihailovich (195–204)*. Bloomington: Slavica.
- Heine, B. (1997). *Cognitive Foundations of Grammar*. New York: Oxford University Press.
- Heine, B. (1993). Conceptual shift: A lexicon of grammaticalization processes in African Languages. *Afrikanistische Arbeitspapiere*, 34/35, 1–322.
- Heine, B., Claudi, U. & Hünnemeyer, F. (1991). *Grammaticalization: A conceptual framework*. Chicago: University of Chicago Press.
- Heine B. & Kuteva, T. (2006). *The Changing Languages of Europe*. Oxford: Oxford University Press.
- Hetzron, R. (1977). 1977. The Gunnan-Gurage languages. *Ricerche* 12. Naples: Istituto Orientale di Napoli.
- Moravcsik, E. A. (1969). Determination. Working Papers on Language Universals. Stanford 1, 64–98.
- Stanković, B. (2023). Types of contexts inducing the grammaticalization of demonstratives into definite articles – the case of a language without articles. Presented at ICHL26, Heidelberg University, September, 2023.

**ГРАМАТИКАЛИЗАЦИЈА НЕОДРЕЂЕНОГ ЧЛАНА
ИНДУКОВАНА КОНТЕКСТОМ *IN STATU NASCENDI***

Бранимир Станковић

Универзитет у Нишу, Филозофски факултет, Ниш, Србија

Резиме

Рад је истраживао проблем постоје ли специфични контексти у којима се дешава граматиализација кардиналног броја „један“ у неодређени члан у српском језику. Досадашња истраживања ове појаве испитивала су може ли се или не број употребити са некардиналном интерпретацијом у одређеном језику, покушавајући да утврде који степен граматиализације је достигнут на петостепеној скали коју је предложио Heine (1997): 1) број; 2) презентативни маркер; 3) маркер специфичности; 4) маркер неспецифичности; 5) генерални члан. Тако, Белај и Матовач (2015) показују да се у хрватском број „један“ може употребити као маркер специфичности, али чак и као део неспецифичне именичке синтагме. Још један од фактора који се разматра у литератури су и језички контакти, за које је недвосмислено

показано да могу да убрзају овај процес, у смислу да језици са неодређеним чланом могу утицати на билингвалне говорнике језика без ове категорије да почну употребљавати број у поменутим окружењима. У нашем раду фокусирали смо се на питање у којим контекстима су говорници српског приморани да употребе ову јединицу како би се добила жељена неодређена (индефините) интерпретација номиналне синтагме, јер би у супротном она имала недвосмислено одређено (дефините) читање. Хипотеза коју смо испитивали гласила је: што је номинални опис у српском језику штурјији/сиромашнији, то је већа вероватноћа да ће говорници бити принуђени да употребе број са некардиналном интерпретацијом да маркирају њену неодређеност (што значи да ће одговарајући примери бити оцењени ниско), и обрнуто – са порастом номиналног описа, опадаће потреба да се обележи неодређеност референта укупног језичког израза (те ће адекватне реченице добити сразмерно више оцене). Сprovedено је онлајн истраживање прагматичке подесности, у коме је 35 изворних говорника српског језика оцењивало прихватљивост циљних примера у датом контексту на скали Ликертовог типа од 1 до 5. Критични примери били су подељени у три групе од по шест стимулуса: 1) „гола“ именица; 2) именица модификована придевом; 3) именица модификована придевом и релативном клаузом или предлошком синтагмом – што је укупно чинило 18 циљних примера, рандомизованих заједно са још 24 дистрактора. Сprovedен је Спирманов тест корелације и добијен коефицијент $r_s = 0.94631$, чиме је недвосмислено потврђена постављена хипотеза. Коначно, добијени резултати сагледани су и у ширем контексту универзалних граматикализацијских процеса, при чему се као логична наметнула и следећа радна хипотеза коју би требало испитати, а то је идеја да су управо одређени типови контекста (сиромашни номинални описи који бивају недвосмислено интерпретирани као одређени) одговорни за покретање граматикализације броја у неодређени члан.