Daniela Katunar
Faculty of Humanities and Social Sciences, Zagreb
dkatunar@ffzg.hr
Jana Willer Gold
Faculty of Humanities and Social Sciences, Zagreb
jwgold@ffzg.hr
Tena Gnjatović
Faculty of Humanities and Social Sciences, Zagreb
tgnjatov@ffzg.hr

Achieving specificity in an articleless language: Specificity markers in Croatian

The main goal of the paper was to detect specificity markers in Croatian. As specificity is in a close relation to (in)definiteness, markers of (in)definiteness such as articles are standardly used as indicators of specificity as well. However, detecting specificity in articleless languages becomes less straightforward. Regardless of the lack of the article system, other (in)definiteness markers exist in Croatian. Here, we focused on the following markers of (in)definiteness in order to find specificity markers: i) the adjectives, ii) the numeral *jedan* 'one', and iii) the demonstrative (definite) and indefinite pronouns. As specificity markers in Croatian we discuss the numeral *jedan* 'one', the indefinite pronoun *neki* 'some' and the construction *taj + neki* 'that some'. Our study also lead to the formation of three categories of specificity based on the properties of these markers: a) contextual specificity, b) semi–lexicalized specificity and c) constructional specificity.

---

1. Introduction

Specificity is a semantic property of noun phrases closely related to the notions of indefiniteness and definiteness, as it refers to the ways of marking noun phrases as expressions with a referring function (see Heusinger 2002a). The particular referring function of specificity is to point out the semantic uniqueness of an entity in the mind of the speaker (Haspelmath 1997, Pavey 2008). In this regard, the dominant characteristics of specificity can be subsumed under the notions of referentiality and identifiability, thus meaning that the entity denoted by the specific NP is a) existentially presupposed and b) the speaker is certain of the identity of the referent (Heusinger 2002a). Dis-

25

cussions related to specificity (e.g. Lyons 1999, Heusinger 2002a, Haspelmath 1997) provide various examples of specificity that illustrate how specificity can be expressed by different lexical or grammatical means, e.g. included within the article system in English, or simply inferred from the context of the utterance itself.

Languages without article systems have been noted to express the notions of indefiniteness/definiteness in various ways, e.g. through the use of pronouns, affixes and so forth (see Haspelmath 1997, Ionin 2010), and thus require a different approach to investigating specificity with regard to the mechanisms employed to express (in)definiteness. Many of these languages are sometimes referred to as 'determinerless' languages (Bošković 2005, 2008, 2009), and some of these, such as Croatian, Serbian and Russian, provide the basis for novel syntactic and semantic analyses of determinacy and (in)definiteness.[1] However, few of these analyses have focused on specificity marking in these languages[2].

What is of interest then is to explore the language specific means that such 'determinerless' languages use to express specificity. Since Croatian is such a language, the goal of this paper is to provide a comprehensive account of the linguistic structures used in expressing specificity in Croatian. As to our knowledge, there have been few studies focusing on specificity marking in Croatian (see Aljović 2002), so we believe that the present paper will make a contribution to the linguistic description of this phenomenon in terms of a) describing the specificity markers in Croatian, b) expanding the descriptions of specificity cross–linguistically, and c) providing new data on which to base an account of specificity and its syntactic, semantic and pragmatic properties.

Since limited data on specificity marking existed for Croatian, the primary task was to find the linguistic elements Croatian employs for marking specificity. For this reason corpus data were consulted[3] and a study of the results

---

1   As we here do not discuss the functional–lexical status of the DP from the perspective of formal/generative approaches, we refer the reader to the following literature. For a general discussion of the topic see Alexiadou, Haegeman and Stavrou 2007; for the discussion of splitting the DP where projections FinP and TopP host [+definite] and [+specific] features respectively see Ihsane & Puskás 2001, Aboh et al. 2010. For the discussion of the topic for Croatian data see two competing proposals: for the proposal of the complete omission of the DP layer in favor of a simple NP analysis see Zlatić 1998, 2009, Trenkić 2004, Bošković 2005, 2008, 2009; for the proposed of elaborated DP structure analysis see Progovac 1998, Leko 1999, Dimitrova–Vulchanova and Giusti 1998, Bašić 2004, Caruso 2011a, 2011b; for the evidence of the D head, hence DP, in Croatian see Caruso 2011a, b. For the discussion of the topic for other Slavic languages see Rutkowski 2002, Pereltsvaig & Franks 2004, Pereltsvaig 2007. We also note that the analysis proposed for the Croatian data in this article is in accordance with the elaborated DP structure analysis, and in particular with the analysis and arguments presented in Caruso 2011a, b.

2   Some exceptions are Heusinger's (2002a) study on specificity markers in Turkish, as well as Ionin's (2006) analysis of specificity with respect to the Russian article *odin* 'one'.

3   Corpora used in our research are the Croatian National Corpus, http://www.hnk.ffzg.hr/, and the Croatian Language Repository, http://riznica.ihjj.hr/. Our analysis of the corpus data is primarily qualitative, focused on describing the main properties of specificity markers. Frequency analysis was applied in some cases in order to confirm our hypotheses about the properties of specificity markers.

was conducted, which yielded the findings of specificity markers as well as their classification according to the morphosyntactic and semantic properties of these markers. The NP constructions most commonly used for expressing specificity in Croatian are a) the constructions with the numeral *jedan* 'one'[4], and b) the Definite + Indefinite Pronoun constructions, e.g. *taj neki* 'that some'.

1.1 Specificity and (in)definiteness

Specificity is often considered as cross–cutting the notions of definiteness and indefiniteness. As Heusinger (2002a) points out, specificity is spread over the two categories in that both specific and non–specific features can combine with indefinite NPs, and the two categories do not yield contradictory interpretations of the NP in question (as shown in example (1)).

(1) I bought <u>a car</u> yesterday.

The NP *a car* gives a specific reading of the NP because the speaker knows the identity of the car and also states that the car is a unique referent, i.e. it is a single specific automobile that is now the property of the speaker. Contextual inference is the predominant factor in the specific reading of the NP, since it is expected that the buyer of the car knows exactly which car he or she is talking about. In this respect, specificity is often discussed as being a subcategory of (in)definiteness, which functions as a fine–grained distinction usually conditioned by the context of the utterance. For this reason many studies conducted on languages with article systems, such as English, focus mostly on specificity as coded within the article system itself (e.g. Fodor and Sag 1982).

It is also often considered that all definite NPs are at the same time specific by default, in that the definiteness of the entity denoted by the NP is a necessary consequence of the identifiability and referentiality of the referent. So for instance, in the example (2), *the blue book* has a specific referent that both the speaker and the hearer recognize, and as such is both definite and necessarily specific.

(2) Give me <u>the blue book</u> on the table over there.

Heusinger (2002a)[5] presents this intersection of the two categories in the following table 1:

identified by	definite (+ specific)	indefinite specific	indefinite non–specific
speaker	+	+	–
hearer	+	–	–

Table 1. Classification of definite and indefinite markers according to their relation with specificity and speaker/hearer knowledge of the referent's identity.

———————————

4 Which is, as will be argued in the paper, also functionally an indefinite article.
5 In his paper Heusinger (2002a) goes on to argue that a fourth category, that of definite non–specific markers, can be established, e.g. *Some day I will find the man of my dreams.*

It is worth noting that the most interesting and the most studied category in the table is the one in the middle, the specific indefinite NPs, mostly because this is the category in which the two notions, that of specificity and indefiniteness are most clearly separated by the speaker/hearer knowledge of the identity of the referent. In our description of the Croatian data, we will also focus on the indefinite – specific markers as the starting point for our analysis.

However, it has been pointed out by Heusinger (2002a) that specificity can be distinguished from indefiniteness based on its focus on "speaker knowledge", whereas indefiniteness in examples such as (1) has its focus on the discourse as a whole. In other words, the function of the article *a* is to mark the novelty of the NP *car* within the discourse in terms of indefiniteness. The indefinite article thus has a pragmatic–discourse function different from specificity, in terms of establishing links between parts of the discourse and marking novel topics of the discourse (see Heusinger 2002a).

Additionally, it has been noted by Ionin (2006) and Heusinger (2002a) that both definite and indefinite NPs can have specific as well as non–specific readings. In her study on the English demonstrative *this*, Ionin (2006) shows how the demonstrative pronoun has a specific indefinite reading when it is used in its unstressed, phonologically reduced form as in example (3):

(3) Mary read this exciting book the other day.

Such distinctions are important to note because they imply that the investigation of specificity in different languages has to take into account possible different functions of specificity and definiteness that could prompt languages to use different linguistic forms for the two categories. More importantly, specificity may be in some cases a property relevant enough to be expressed formally, rather than only through context. Many languages can draw from existing definiteness markers to express specificity as well (such is the case of English, for example).

In his analysis on the ways of expressing specificity in various languages, Heusinger (2002b) calls these linguistic forms *language–specific grammatical means of expressing specificity* and includes primarily grammatical constructions that exploit the existing (in)definiteness markers by combining them into specificity markers.

The clearest example of specificity as a standalone category comes from languages such as Maori or Samoan that have particles that function as specificity markers instead of (in)definiteness markers. In a study of Samoan conducted by Mosel and Hovdhaugen (1992) (in: Lyons 1999) the distributional patterns of two Samoan particles *le* and *se* lead the authors to conclude that these are specificity particles marking 'specific' and 'non–specific' NPs, respectively. The example they provide in (4) illustrates how these distributional properties guide the usage of the 'specific particle' *le* regardless of the discourse functions performed by the indefinite and definite articles in its English translation.

28

(4) Sa   i ai **le**   ulugali'i 'o Papa **le** tane   a   'o
    PST exist ART couple   PRS Papa ART husband but PRS
    Eleele **le** fafine.
    Eleele ART woman
    'There was **a** couple, Papa, **the** husband, and Eleele, **the** wife.'

As Lyons (1999) points out, the distributional properties of the Samoan particles depend only on the specificity or non–specificity of the NPs, and not on the novelty of NPs as topics in the discourse. In other words, the particles have a referring function, and not a discourse pragmatic one. In the example it is clear how the knowledge of the narrator about a particular couple determines the usage of the specific particle, and not the fact that the NP 'a couple' is introduced for the first time into the discourse. The English translation shows how different principles guide the usage of its articles as opposed to Samoan. Similarly, the Siouan language Lakhota distinguishes the indefinite specific article and the indefinite non–specific article (Van Valin 1995: 41–42).

Although languages such as Samoan and Lakhota provide clear evidence of the existence of specificity as a linguistic category in its own right, many languages that fall between the examples of Samoan and English still remain to be investigated. Such is the case of Croatian, as we will argue below.

2. Indefiniteness in Croatian

To examine the relation between (in)definiteness and specificity in Croatian, we must first examine the ways of expressing definiteness in Croatian. Croatian does not have an article system, so definiteness is often inferred from the context of the utterance, e.g.

(5) *Ptica   je   odletjela sa   stabla.*
    bird–NOM.SG AUX–3.SG fly–PST   from tree–GEN.SG
    'A / The bird flew off the tree.'

The most common means of explicitly expressing definiteness according to Croatian grammars (Silić and Pranjković 2007; Barić et al. 2005) are a) via indefinite and definite adjective inflection, b) definite and indefinite pronouns, and c) the numeral *jedan* 'one' in the function of an indefinite article.

2.1 The definite and indefinite adjectival inflection

Adjectives in the attributive position in Croatian, as in the majority of Slavic languages, can appear in two different forms – the definite and the indefinite form (e.g. Aljović 2002, Silić i Pranjković 2007: 134), also referred to as the long and the short form of the adjective, or the adjectival aspect (Katičić 1991: 385, Znika 2006: 16). For example, the nominative masculine singular of the adjective meaning 'big' is *velik* if indefinite, and *veliki* if definite. It is usually claimed, as in Silić and Pranjković (2007), that the indefinite form of the adjectives is used to modify a noun phrase that is discourse–new and unknown

to the speech act participants, while definite adjectives function as modifiers in NPs that are known and have been previously mentioned in the discourse.

The distinction between the definite and the indefinite adjectival inflection is not so clear–cut in language use anymore. Silić and Pranjković (2007: 381) state that the lack of this distinction is characteristic of the language of administration, and Brozović (1988, as cited in Znika 2006) believes that the opposition has become optional in language and is stylistically marked. Any study of the meaning and use of long as opposed to short adjectival forms in the attributive position should thus take into account the neutralization of this distinction in contemporary language.

Examples that we found in the Croatian Language Repository confirm the claim that both long and short adjectival forms can be used equally in definite and indefinite noun phrases, as shown in (6) and (7):

(6) Taj   velik   skup,   na kojem   će   *sudjelovati*
that[NOM.SG] big[INDEF.NOM.SG] assembly[NOM.SG] on which–LOC.SG AUX–3SG participate–INF
stotinjak likovnih kritičara   iz   čitavoga   svijeta,
hundred   art–GEN.PL critic–GEN.PL from entire–GEN.SG world–GEN.SG
a pratit će ga izložbe u Zagrebu i Dubrovniku, najavili su u petak u Muzeju »Mimara« predsjednik hrvatske sekcije AICA–e Berislav Valušek i tajnica Leonida Kovač.
'This big assembly, which around a hundred art critics from all around the world will take part in, and which will be accompanied by exhibitions in Zagreb and Dubrovnik, was announced on Friday in the "Mimara" Museum by the president of the Croatian section of AICA, Berislav Valušek, and the secretary, Leonida Kovač.'

(7) Svi se rađamo, svi živimo i umiremo, ali neka umiremo *tako, kako je rekao*
as AUX–3SG   say–PST
jedan   *veliki*   *pjesnik, da svijet iza sebe ostavimo*
one[NOM.SG] big–DEF.NOM.SG poet[NOM.SG]
malo boljim, nego što smo ga našli.
'We are all born, we all live and die, but let us die in such a way, as a great poet said, that we leave the world behind a bit better than we found it.'

In (6), an indefinite adjective occurs in an NP whose definiteness is signaled by the demonstrative pronoun *taj*, while in (7) the numeral / article *jedan* 'a; one' indicates the indefiniteness of the NP (see Section 2.3. below) in which, nevertheless, a "definite" adjective is used as modification. This clearly shows that the distinction between long and short adjectival forms, if it ever did carry the definite/indefinite meaning, should be reinterpreted as to how it functions in the contemporary language, possibly as a specificity marker.

Aljović (30–31) uses very similar examples to show that the two adjectival inflections are not used to mark the definiteness/indefiniteness distinction, but rather to trigger a specificity reading. In her view, the short form of the adjective is always used in non–specific NPs, while the long form signals that

the noun phrase should be interpreted as specific, and not necessarily definite, as, for example, in (8):

(8) *jedan vrijedni student*
    'one diligent student'

While we agree with Aljović that long and short adjectival forms are not used to signal the definiteness or indefiniteness of the NP, in this case we are more inclined to attribute the specificity reading of (8) to the numeral / article *jedan*, 'a; one' which we believe to be an instance of semi–lexicalized specificity (see section 3.).

2.2 Indefinite and definite pronouns

As to the pronoun system, Croatian grammars often classify indefinite and definite pronouns as specific subtypes of the pronoun word class. Definite pronouns are termed demonstratives, and their primary function is defined as indexical, pointing to the proximity of a certain entity to the speaker, hearer or a third party (Barić et al. 2005).

(9) *Evo ti ova knjiga, ti meni daj tu,*
    here 2SG–DAT DEM.PROX–NOM.SG book–NOM.SG 2SG.NOM 1SG–DAT give–IMP.2SG DEM.MED–ACC.SG
    a ona na stolu mi ne treba.
    and DEM.DIST–NOM.SG on table–LOC.SG 1SG–DAT NEG need–PRS.3SG
    'Take this book, you give me that one, and that one on the table I do not need.'

It is also important to note that this three–way distinction is slowly being replaced by the dyadic opposition *ovaj – taj* 'this–that', especially in spoken language and substandard registers[6].

The function of demonstratives to express definiteness is not uncommon among languages without an article system, as a way of compensating for the communicative needs of expressing definiteness (see Lyons 1999, Haspelmath 1997). In this way Croatian demonstratives follow the pattern observed in other languages, e.g. Russian, Czech, etc. (Haspelmath 1997)[7]. Furthermore, it can be shown that examples such as:

(10) Svi smo čuli tu priču *nekoliko puta.*
    everyone–NOM.PL AUX–1PL hear–PST DEM.MED–ACC.SG story–ACC.SG several times
    'We have all heard the / that story several times.'

with the demonstrative *tu* 'that', can be translated felicitously both with the definite article *the* or the demonstrative *this/that.*

Indefinite pronouns form a large subclass of pronouns with both simplex and complex forms, e.g. *tko* 'who, someone' > *ne–tko* 'someone, somebody', *ni–tko* 'no one, nobody', *i–tko* 'anyone, anybody', *sva–tko* 'everyone, everybody',

---

6   See also Žic Fuchs (1996) for a discussion on spatial deixis in Croatian.
7   This is a very common definiteness marking strategy is languages without definite articles, demonstratives are used in the same way in e.g. Mandarin Chinese (Hedberg 1996).

etc. Their main function is to express indefiniteness (Silić and Pranjković 2000; Barić et al. 2005). Among indefinite pronouns one has to distinguish between those that take the modifier position in the NP[8], e.g. *neki* 'some', *nekakav* 'some kind' and *nečiji* 'someone's', and those that take the head position, e.g. *netko* 'someone, somebody', as in examples (11) and (12):

(11) Neki   čovjek   me   pozdravio u dućanu.
    some–NOM.SG man[NOM.SG] 1SG–ACC greet–PST   in store–LOC.SG
    'Some man said hello to me in the supermarket.'

(12) Netko   me   pozdravio u dućanu.
    someone–NOM.SG 1SG–ACC greet–PST   in store–LOC.SG
    'Somebody said hello to me in the supermarket.'

In the example (11) the pronoun *neki* 'some' establishes the indefiniteness of the NP čovjek 'man' and thus acts as an indefiniteness marker.

As to the correlation between the demonstrative (definite) and indefinite modifier pronouns, it can be stated that they perform the discourse–pragmatic functions in a very similar way to articles in English, e.g. the correlation between the pronouns *neki* 'some' and *taj* 'that' in example (13):

(13) <u>*Neki*</u>   me   čovjek   pozdravio u dućanu   ujutro.
    some–NOM.SG 1SG–ACC man[NOM.SG] greet–PST   in store–LOC.SG in.the.morning
    *Mislim   da bi   <u>*taj*</u>   čovjek   mogao   biti*
    think–PRS.1SG that AUX–3SG DEM.MED–NOM.SG man[NOM.SG] can–COND be–INF
    tvoj   brat.
    your[NOM.SG] brother[NOM.SG]
    '<u>A / some</u> man said hello to me in the supermarket this morning. I think <u>the / that</u> man might be your brother.'

2.3 The article jedan

Although Croatian lacks articles as a separate word class, the numeral *jedan* 'one' in some cases acts like an indefinite article (see Silić and Pranjković 2007), e.g.

(14) Tu bi   sad dobro došlo   jedno   majčino   srce.
    there AUX–3SG   now well   come–COND one–NOM.SG mother's–NOM.SG heart–NOM.SG
    'A mother's heart would be of use right now.'

(15) Što   može   jedna   žena   učiniti suočena s
    what–ACC can–PRS.3SG one–NOM.SG woman–NOM.SG do–INF face–PTCP with
    takvom   *tragedijom?*
    such–INS.SG tragedy–INS.SG
    'What can a woman do in face of such tragedy?'

---

8   In this paper we will only focus on indefinite pronouns as modifiers, since they are the ones that can carry an "article–like" function in Croatian, contrary to indefinite pronouns in the head position, e.g. *netko* 'somebody'.

According to Silić and Pranjković (2007), the use of *jedan* 'one' as an article is restricted only to cases where the adjective of the NP is not inflectionally marked as either a definite or an indefinite adjective, e.g. *majčino srce* 'mother's heart'. However, due to the fact that definite and indefinite adjectival flection is becoming obsolete, the use of *jedan* 'one' as an article[9] seems to be more facultative and influenced by specificity marking as well (see discussion below). As the following corpus examples[10] show, *jedan* 'one' can be used as an article either with a definite or an indefinite adjective:

(16) ...napisao bi   u ovom   pismu   da baš čita
     write–COND AUX–3SG in this–LOC.SG letter–LOC.SG that just read–PRS.3SG
     jedan novi   hrvatski   roman...
     one[ACC.SG] new–DEF.ACC.SG Croatian–ACC.SG novel–ACC.SG
     '...he would write in this letter that he is just reading a new Croatian novel...'

(17) *Iz njih   struji   **jedan nov**   senzibilitet kojeg*
     from 3PL–GEN flow–PRS.3SG one[NOM.SG] new[INDEF.NOM.SG] sensibility[NOM.SG] which–ACC.SG
     bismo mogli   nazvati mješavinom buntovništva i   romantike.
     AUX–1PL can–COND call–INF   mixture–INS.SG rebellion–GEN.SG and romance–GEN.SG
     'Out of them flows a new sensibility which we could call a mixture of rebellion and romanticism.'

3. Specificity markers in Croatian

Though there is much more that can be said about definiteness marking in Croatian, for the purposes of our research we decided to focus on the aforementioned strategies of marking (in)definiteness to determine the factors that guide specificity readings in Croatian. Based on the fact that markers of specificity are usually closely tied to those of (in)definiteness, our corpus search focused on these markers and their syntagmatic surroundings within a discourse. To determine the specificity markers we used the criterion of substitution, according to which specificity markers are substituted in a sentence with 'a certain' (see Heusinger 2002a, Lyons 1999). This simple test allowed us to establish two prominent specificity markers, the article *jedan* 'one' and the construction *taj+neki* 'that+some', i.e. Definite + Indefinite Pronoun construction.

---

9   Use of the numeral 'one' as an indefinite article is not uncommon among (especially Slavic) languages, see Ionin (2006) for the description of the Russian *odin* 'one'. Diachronically, many languages have developed articles by grammaticalizing numerals, e.g. French and English (see Lyons 1999). Turkish seems to be very similar to Croatian in this respect, since it uses *bir* 'one' to mark specific indefinite NPs in the accusative case (Schaaik 1996: 40–41).

10   http://riznica.ihjj.hr/.

(18) **Jedan**   se   zadarski   trgovac   potkraj   toga
one[NOM.SG] REFL Zadar–NOM.SG merchant[NOM.SG] at.the.end DEM.MED–GEN.SG
stoljeća   *može*   *gorditi*   *posjedovanjem rukopisnoga*
century–GEN.SG can–PRS.3SG be.proud–INF owning–INS.SG   manuscript–GEN.SG
primjerka   cijele   Božanstvene   komedije.
copy–GEN.SG entire–GEN.SG Divine–GEN.SG comedy–GEN.SG
'A certain merchant from Zadar by the end of the century can be proud
of owning a manuscript example of the entire Divine Comedy.'

(19) *Larry*   *prvo mora*   *odraditi* **taj**   **neki**   *susret.*
Larry–NOM.SG first   must–PRS.3SG do–INF   DEM.MED–ACC.SG some–ACC.SG meeting[ACC.SG]
'Larry first has to go to a certain meeting.'

Based on the corpus overview we also established a classification of the ways specificity is expressed in Croatian. The classification was based on the following factors influencing specificity readings:

a) because of contextual inference, indefinite articles and pronouns can have both specific and non–specific readings
b) some linguistic expressions yield only specific readings, and for that reason must be analyzed as having a systematic specificity function, not simply being subjected to contextual modification
c) a specific NP has to be a) identifiable and b) carry felicity conditions on its use (see Ionin 2010)

The contrast between contextual modification and systematic meaning is the key opposition in determining specificity markers, i.e. the language specific means Croatian uses to express specificity[11]. The felicity conditions of specific NPs are discussed at length by Ionin (2010) and refer to the fact that certain pragmatic conditions have to be met for the utterance to sound perfectly acceptable. For instance, in the following example (20) the use of *this* would be odd if the speaker did not know anything about the movie (genre, title or the like), as compared to (21):

(20) ?Mary wants to see <u>this new movie;</u> I don't know which movie it is.
(21) Mary wants to see some new movie; I don't know which movie it is.

---

11   Within the framework of usage–based models of language (Kemmer and Barlow 2000, Evans and Green 2004), the term contextual modification refers to meanings of lexical units that are established through the surrounding context of the utterance and are not part of the encyclopedic knowledge pertaining to the language system, more specifically the conceptual structure of the lexical unit itself. Evans and Green (2004) give examples such as *This toy is safe for children* vs *The girl is safe with her grandmother,* where the meaning of *safe* is in both cases ruled by the surrounding context. In other words, such modifications are pragmatically guided occurrences, which depend on knowledge structures wider than the lexical meaning itself. Such is the example in (1) 'I bought a car yesterday', where the specificity reading is a matter of contextual modification. Some frequent instances of contextual modifications can become more and more entrenched in time, and this process can be observed as a tendency of some language units to appear more frequently with the same meaning; such is the process between the numeral *jedan* 'one' and the grammaticalized article *jedan* 'a, one'.

34

The classification we propose based on these factors deals with three categories of specificity:

1) contextual specificity – zero marked, indefinite pronoun *neki*
2) semi–lexicalized specificity – article *jedan*
3) constructional specificity – *taj+neki* construction

3.1 Contextual specificity

The first category, contextual specificity, is the most straightforward pragmatically motivated way of establishing specificity. In Croatian it is encountered in examples such as:

(22) *Želim   se   oženiti **0** – Norvežankom.*
    want–PRS.1SG REFL marry–INF   Norwegian–INS.SG
    'I want to marry a Norwegian.'

The example lacks any prenominal (in)definiteness markers and can have both a specific and a non–specific reading depending on the context which can be expanded as follows:

(23) *Želim   se   oženiti **0** – Norvežankom koju   sam   upoznao*
    want–PRS.1SG REFL marry–INF   Norwegian–INS.SG who–ACC.SG AUX–1SG meet–PST
    prošlog   ljeta   u Dubrovniku.
    last–GEN.SG summer–GEN.SG in Dubrovnik–LOC.SG
    'I want to marry a/the Norwegian that I met last summer in Dubrovnik.' – specific

(24) *Želim   se   oženiti **0** – Norvežankom, ali još nisam*
    want–PRS.1SG REFL marry–INF   Norwegian–INS.SG but yet AUX.NEG–1SG
    upoznao nijednu.
    meet–PST none–ACC.SG
    'I want to marry a Norwegian but I haven't met one yet.' – non specific

When the indefinite pronoun *neki* 'some' is added to the existing sentences they are still both grammatically acceptable, however the example (25) seems pragmatically odd.

(25) ?*Želim   se   oženiti **nekom**   Norvežankom koju*
    want–PRS.1SG REFL marry–INF some–INS.SG Norwegian–INS.SG who–ACC.SG
    sam   upoznao prošlog   ljeta   u Dubrovniku.
    AUX–1SG meet–PST last–GEN.SG summer–GEN.SG in Dubrovnik–LOC.SG
    'I want to marry a Norwegian that I met last summer in Dubrovnik.' – specific

(26) *Želim   se   oženiti **nekom**   Norvežankom, ali još*
    want–PRS.1SG REFL marry–INF some–INS.SG Norwegian–INS.SG but yet
    nisam   upoznao nijednu.
    AUX.NEG–1SG meet–PST none–ACC.SG
    'I want to marry a Norwegian but I haven't met one yet.' – non specific

35

The differences in co–reference (*neka – koja / neka – nijedna*) between the two sentences are often referred to as scope properties of specific NPs (see Pavey 2008, Fodor and Sag 1982, Heusinger 2002a, Lyons 1999). As Ionin (2010) points out, specific NPs have wide scope properties, meaning that the specific NPs remain identifiable as a constant through the utterance and can co–reference as such with other elements within the utterance. In an example taken from Lyons (1999) such is the scope relation between the NP 'a merchant banker' and the co–referenced pronoun 'her':

(27) Peter intends to marry <u>a merchant banker</u> – even though he doesn't get on at all with <u>her</u>.

Non–specific NPs, on the other hand, have different scope properties since the non–specific NP is a variable that cannot be co–referenced in the same way:

(28) Peter intends to marry <u>a merchant banker</u> – though he hasn't met <u>one</u> yet.

Thus the scope properties are also a good indicator of specificity, and can point to specificity markers alongside the so–called substitution test with 'a certain'.

Although we list the indefinite pronoun as an example of contextual specificity, it seems that without postmodification it tends to yield mostly non–specific reading (see e.g. (33) and (34)). However, the frequency analysis that we have performed to establish its tendency to lexicalize non–specific reading showed that the distinction between the specific and non–specific *neki* is approximately only 52% for the non–specific reading. We believe that the reason why the corpus data is inconclusive may be attributed to the properties of corpus texts in general, which consist mostly of complex sentences including postmodifications of the pronominal construction such as a relative clause in example (29).

(29) *Neki   čovjek   koji se   spremao   na dugo*
some–NOM.SG man[NOM.SG] who–NOM.SG REFL prepare–PST on long.ACC.SG
putovanje, razmišljao je kako da smanji trošak.[12]
journey–ACC.SG
'A man who was preparing himself for a long journey, was thinking about how to cut down on his expenses.'

3.2 Semi–lexicalized specificity

If we add the article *jedan* 'a; one' to our examples, we will get a wide scope specific NP:

(30) *Želim   se   oženiti   **jednom** Norvežankom   koju*
want–PRS.1SG REFL marry–INF one–INS.SG Norwegian–INS.SG who–ACC.SG
sam   upoznao prošlog   ljeta   u Dubrovniku.
AUX–1SG meet–PST last–GEN.SG summer–GEN.SG in Dubrovnik–LOC.SG
'I want to marry a Norwegian that I met last summer in Dubrovnik.'
– specific

---

12   http://riznica.ihjj.hr

On the other hand, although the examples such as (31) are grammatically acceptable, they seem pragmatically odd:

(31) ?*Želim   se   oženiti   **jednom** Norvežankom,   ali*
        want–PRS.1SG REFL marry–INF one–INS.SG Norwegian–INS.SG but
        još nisam   upoznao nijednu.
        yet AUX.NEG–1SG meet–PST none–ACC.SG
        'I want to marry a Norwegian, but I haven't met one yet.'

Furthermore, if we contrast examples with *jedan* 'a; one' and non–marked or pronoun marked indefinite NPs, there is a clear contrast in the specific versus non–specific interpretation of these NPs, as in the examples:

(32) *Selim   se   u 0 – ljepši grad.* Indefinite specific
        move–PRS.1SG REFL in nicer–ACC.SG city[ACC.SG]
        'I'm moving to a nicer city.'

(33) *Selim   se   u **neki**   ljepši   grad.* Indefinite non–specific
        move–PRS.1SG REFL in some–ACC.SG nicer–ACC.SG city[ACC.SG]
        'I'm moving to some nicer city.'

Note the oddness of the example:

(34) *?Selim   se   u **neki**   ljepši   grad.*
        move–PRS.1SG REFL in some–ACC.SG nicer–ACC.SG city[ACC.SG]
        Taj   je   grad   Zagreb. Indefinite non–specific
        DEM.MED–NOM.SG COP–3SG city[NOM.SG] Zagreb[ACC.SG]
        ?'I'm moving to some nicer city. It is the city of Zagreb.'

However, the following is fully acceptable:

(35) *Selim   se   u **jedan**   ljepši   grad.*
        move–PRS.1SG REFL in one–ACC.SG nicer–ACC.SG city[ACC.SG]
        Taj   je   grad   Zagreb. Indefinite specific
        DEM.MED–NOM.SG COP–3SG city[NOM.SG] Zagreb[ACC.SG]
        'I'm moving to a / one nicer city. It is the city of Zagreb.'

Thus it seems that *jedan* 'a; one' is used predominantly as a specific indefinite marker, and instantiates the category of lexicalized specificity in Croatian. This is not to say that *jedan* 'a; one' has to be a specific marker in every instance of it's usage, but only that it shows a strong tendency towards specific reading. Based on the opposition between contextual modification and systematic meaning, it serves to show that the lexicalization of specificity in the example of *jedan* 'a; one' is a gradable phenomenon. However, it is a phenomenon worth noting since it expands the description of the use of *jedan* 'a; one' beyond the current descriptions. On the other hand, the same can be said for the indefinite pronoun *neki* 'some', which shows a tendency towards indefinite non–specific readings when the NP is not postmodified, as in exam-

ple (32). To further support this claim, we performed a frequency analysis[13] of the specific versus non–specific occurrences of *jedan* 'a; one' in the Croatian National Corpus (henceforth CNC).

The analysis showed the following results:

jedan	tokens: numeral and article	numeral	article	article specific	article non–specific
Absolute frequency	197973	160358	37614	31677	5937
Relative frequency	0,197%	0,16%	0,037%	0,031%	0,0059%

Table 1. Frequency analysis of the occurrences of *jedan* 'a; one' in the CNC.

Column 1 represents the sum of all the tokens of *jedan* in the CNC, columns 2 and 3 represent the token number of numeral *jedan* 'one' and the article *jedan* 'a; one', respectively. Columns 4 and 5 show the frequencies of specific vs non–specific occurrences of the article *jedan* 'a; one'.

3.3 Constructional specificity

The category of constructional specificity is perhaps the most clear–cut case of language specific means Croatian uses to express specificity. In all the corpus examples of the construction *taj+neki* 'that+some' the readings were specific[14], i.e. had a wide scope and were substitutable with 'a certain'[15]. It is interesting to note that the construction itself is composed of two notionally opposing parts, the demonstrative pronoun *taj* 'that' and the indefinite pronoun *neki* 'some'. Although the two units of the construction have opposing semantic features, those of definiteness and indefiniteness, it seems that the end result is a construction with the features 'Indefinite and Specific'. From the examples it is clear that the criterion of speaker knowledge is very prominent, e.g.

(36) *Prije nego sam otišla u London znala sam* **ta**
know–PST AUX–1SG DEM.MED–ACC.PL

neka *nova estradna imena.*
some–ACC.PL new–ACC.PL scene–ACC.PL name–ACC.PL

'Before I went to London, I knew these new names on the music scene.'

---

13 The Croatian National Corpus has 100 million tokens, and it is compiled from various literary and newspaper texts. The analysis was conducted on a random sample of 200 concordances, by determing the numeral, article, and specific and non–specific article readings of *jedan,* and then multiplying these percentages with the total number of tokens of *jedan,* thus obtaining an estimation of the distribution of various meanings of *jedan.*

14 All the inflectional variants for gender and number were examined, e.g. *ta neka, ti neki, te neke, taj neki.*

15 Also note that the order of the elements is reverse than that of the English 'a certain' construction, which has the indefinite marker first.

(37) *Iz tog nekog jako strukturiranog prostora Globe–a,*
from this–GEN.SG some–GEN.SG very structured–GEN.SG space–GEN.SG Globe–GEN.SG
ovog se ljeta seli na cijeli otok.
this–GEN.SG REFL summer–GEN.SG move–PRS.1SG on entire–ACC.SG island.ACC.SG
'From this very structured space of the Globe, this summer it is moving to the entire island.'

In both examples taken from the CNC it is implied that the speaker knows the identity of 'new names on the music scene' and 'the structured space', respectively. Thus it seems that the construction Def. + Indef. pronoun serves a particular function within the discourse which adds the notion of speaker knowledge as a relevant notion of the utterance. In her study on specificity marking in English and Russian, Ionin (2006) called indefinite specific markers, such as English *this* and Russian *odin* 'one' markers of epistemic specificity. The function of such markers, according to Ionin (2006), is to stress the importance of the identity of the indefinite NP, as opposed to cases with *a* articles, where the identity is not of importance. To illustrate this point, Ionin used the following examples for English, where the identity of the referent is necessary for a felicitous interpretation of the sentence (38) and (39).

(38) I want to see this new movie tonight. – identity known

(39) I want to see a/*this new movie tonight. – identity unknown

The Croatian examples follow the felicity conditions established by Ionin (2006) for English:

(40) Želim pogledati taj neki novi
want–PRS.1SG watch–INF DEM.MED–ACC.SG some–ACC.SG new–DEF.ACC.SG
film navečer.
movie–ACC.SG tonight
'I want to see this new movie tonight' – identity known

(41) Želim pogledati *taj neki novi
want–PRS.1SG watch–INF DEM.MED–ACC.SG some–ACC.SG new–DEF.ACC.SG
film navečer.
movie–ACC.SG tonight.
'I want to see this new movie tonight.' – identity unknown

Also, the examples below show the non–specific readings of *neki* 'some'.

(42) Želim pogledati neki novi film navečer.
want–PRS.1SG watch–INF some–ACC.SG new–DEF.ACC.SG movie–ACC.SG tonight
'I want to see some new movie tonight.' – identity unknown

(43) Želim pogledati *neki novi film navečer.
want–PRS.1SG watch–INF some–ACC.SG new–DEF.ACC.SG movie–ACC.SG tonight
'I want to see some new movie tonight.' – identity known

Based on the tendency of *neki* 'some' to have a non–specific reading in many non–postmodified NPs, it seems that the demonstrative *taj* serves to

39

modify this reading into a specific one, thus creating a construction with novel specificity properties. The whole construction being indefinite, we can assume that the original indexical function of *taj* 'that' is of greater salience in creating the construction than its function as a definiteness marker. Thus it loses its definiteness meaning in the *taj neki* 'that some' construction and triggers a specific interpretation of the referent of the NP, creating a specific NP.

4. Conclusion

Specificity is a category that often cross–cuts the category of (in)definiteness. In languages such as Croatian, where no article system exists, these categories have various ways of being expressed and serving their communicative functions. Distinguishing specificity as opposed to (in)definiteness is not an easy task, since the categories often use the same or somewhat related markers. However, since specificity is connected mainly with the speaker knowledge of the identity of the referent, it obviously serves a function that requires its expression in certain linguistic forms, such as the numeral *jedan* 'a; one' and *taj neki* 'that some' in Croatian. Specificity is often a matter of contextual inference, and as such carries various pragmatic information about an utterance. Because of its pragmatic nature, it is often difficult to establish clear–cut boundaries between specific and non–specific markers, as we have shown with the example of *jedan* 'a; one'. Although there are, in Heusinger's terms, *language–specific means* of expressing specificity, we believe that the phenomenon of specificity as a linguistic category is gradable, and for that reason we propose a threefold classification of specificity in Croatian, a) contextual specificity, b) semi–lexicalized specificity, and c) constructional specificity. In our classification we focused on single lexical units or constructions, which can serve as specificity markers, but also showed that the only category with a continuous specificity reading in all cases is the construction *taj neki* 'that some'.

LIST OF ABBREVIATIONS

1 – first person
2 – second person
3 – third person
ACC – accusative
ART – article
AUX – auxiliary
COND – conditional
COP – copula
DAT – dative
DEF – definite
DEM – demonstrative
DIST – distal
GEN – genitive
IMP – imperative

40

INDEF – indefinite
INF – infinitive
INS – instrumental
LOC – locative
MED – medial
NEG – negation
NOM – nominative
PL – plural
PROX – proximal
PRS – present
PST – past
PTCP – participle
REFL – reflexive
SG – singular

References

Aboh, E. O., M. Dyakonova, M. van Koppen (2010) "DP–internal information structure: Some introductory remarks". *Lingua* 120: 782–801. Special issue on DP–internal information structure

Alexiadou A., L. Haegeman, M. Stavrou (2007) *Noun phrase in the generative perspective*. Berlin: Mouton de Gruyter.

Aljović, N. (2002) "Long Adjectival Inflection and Specificity in Serbo–Croatian". *Recherches Linguistiques de Vincennes* 31: 27–42.

Barić et al. (2005) *Hrvatska gramatika* [Croatian Grammar]; Školska knjiga, Zagreb.

Bašić, M. (2004) *Nominal Subextractions and the Structure of NPs in Serbian and English*. MA Thesis: University of Tromsø.

Bošković, Ž. (2005) "On the locality of left branch extraction and the structure of NP". *Studia Linguistica* 59: 1–45.

Bošković, Ž. (2008) "What will you have, DP or NP?", *Proceedings of NELS* 37.

Bošković, Ž. (2009) "More on the no–DP analysis of articleless languages". *Studia Linguistica* 63 (2): 187–203.

Caruso, Đ. Ž. (2011a) "In support of a DP–Analysis of Nominal Phrases in Croatian". *Workshop on Languages With and Without Articles*, Université de Paris 8 & UMR 7023, March 3–4, 2011.

Caruso, Đ. Ž. (2011b) "Nominal Phrases in Croatian as DPs". *Online Proceedings of GLOW in Asia Workshop for Young Scholars* 2011.

Dimitrova–Vulchanova, M., G. Giusti (1998) "Fragments of Balkan Nominal Structure". In: A. Alexiadou and C. Wilder (eds). *Possessors, Predicates and Movement in the Determiner Phrase*. 333–60. Amsterdam/Philadelphia: John Benjamins.

Evans, V., M. Green (2004) *Cognitive Linguistics: An Introduction*. Cambridge University Press: UK

Fodor, J. D., I. Sag (1982) "Referential and quantificational indefinites". *Linguistics and Philosophy* 5: 355–398.

Haspelmath, M., (1997) *Indefinite Pronouns*. Oxford Studies in Typology and Linguistic Theory. Oxford: Clarendon Press.

Hedberg, N. (1996) "Word Order and Cognitive Status in Mandarin Discourse". In: Fretheim, T., J. Gundel (eds). *Reference and Referent Accessibility*. *Pragmatics and Beyond* Series. John Benjamins: 173–192.

von Heusinger, K. (2002a) "Specificity and Definiteness in Sentence and Discourse Structure". *Journal of Semantics* 19: 245–274.

von Heusinger, K. (2002b) "The Cross–Linguistic Implementations of Specificity". In: Jaszczolt K., K. Turner (eds.). *Meaning through Language Contrast*. Vol. 2. Amsterdam; Philadelphia: Benjamins: 405–421. (Pragmatics and Beyond NS 100)

Ihsane, T., G. Puskás (2001) "Specific is not Definite". *Generative Grammar in Geneva* 2: 39–54.

Ionin, T. (2006) "This is definitely specific: specificity and definiteness in article systems". *Natural Language Semantics* 14: 175–234.

Ionin, T. (2010) "The scope of indefinites: an experimental investigation". *Natural Language Semantics* 18.3: 295–350.

Katičić, R. (1991) *Sintaksa hrvatskoga književnog jezika: nacrt za gramatiku*. Zagreb: HAZU and Globus.

Kemmer, S., M. Barlow (eds.) (2000) *Usage–Based Models of Language*. Cambridge University Press.

Leko, N. (1999) "Functional Categories and the Structure of the DP in Bosnian". In Dimitrova–Vulchanova, M., L. Hellan (eds): *Topics in South Slavic Syntax and Semantics*. John Benjamins Publishing Company, Amsterdam: 229–252.

Lyons, Ch., (1999) *Definiteness*. Cambridge University Press.

Pavey, E. (2008) "Predication and reference: the interpretation of noun phrases in specificational constructions". In: Van Valin, R. Jr. (ed) *Investigations of the Syntax–Semantics–Pragmatics Interface*. John Benjamins. 305–317.

Progovac, L. 1998. "Determiner Phrase in a language without determiners". *Journal of Linguistics* 34(1): 165–179.

Pereltsvaig, A. (2007) "On the Universality of DP: A View from Russian". *Studia Linguistica*. 61(1): 59–94.

Pereltsvaig, A., S. Franks (2004) "Functional Categories in the Nominal Domain". In *Proceedings of FASL* 12. The Ottawa Meeting 2003, ed. by Arnaudova, Olga, Wayles Browne, Maria Luisa Rivero, and Danijela Stojanovic. University of Michigan Press. Ann Arbor, MI: 109–128.

Rutkowski, P. (2002) "Noun/Pronoun Asymmetries: Evidence in Support of the DP Hypothesis in Polish". *Jezikoslovlje* 3.1–2: 159–170.

van Schaaik, G. (1996) *Studies in Turkish Grammar*. Wiesbaden: Harrassowitz.

Silić, J., I. Pranjković (2007) *Gramatika hrvatskoga jezika: za gimnazije i visoka učilišta*. (Second edition) Zagreb: Školska knjiga.

Trenkić, D. (2004) "Definiteness in Serbian/Croatian/Bosnian and some implications for the general structure of the nominal phrase". *Lingua* 114: 1401–1427.

Van Valin Jr., R. D. (1995) "Toward a functionalist account of so–called 'extraction constraints'". In: B. Devriendt (ed.) *Complex structures: A functionalist perspective*. Berlin: Mouton de Gruyter: 29–60.

Zlatić, L. (1998) "Slavic Noun Phrases are NPs not DPs". To appear in Harves, S., J. Lavine (eds) (2013): *Comparative Slavic Morphosyntax*. Slavica Publishers, Bloomington.

Zlatić, L. (2009) "Definiteness and Structure of NPs in Slavic". *The Eighth European Conference on Formal Description of Slavic Languages*, December 2–5, 2009, Potsdam, Germany.

Znika, M. (2006) "Kategorija odredenosti i predikatno ime". *Jezik: časopis za kulturu hrvatskoga književnog jezika* 53, 1; 16–25.

Žic Fuchs, M. (1996) "Here" and "There" in Croatian: a case study of an urban standard variety"; in: Dirven, R. (ed.) (1996) *Construal of Space in Language and Thought*. Walter de Gruyter.

42

Izražavanje specifičnosti u jeziku bez članova: Pokazatelji specifičnosti u hrvatskom jeziku

Glavni cilj ovoga rada bio je odrediti pokazatelje specifičnosti u hrvatskom jeziku. Definicija specifičnosti odnosi se na mogućnost govornikove identifikacije jedinstvenog referenta imenske fraze i na taj način funkcionira kao primarno pragmatički utemeljena kategorija, koja svoju funkciju ostvaruje u govornom, odnosno komunikacijskom činu. Također, specifičnost se kao sintaktičko–semantička kategorija u literaturi često veže uz pojmove određenosti i neodređenosti te se obično izražava članovima u jezicima poput engleskoga. Na taj se način specifičnost u literaturi definira kao kategorija koja presijeca kategorije određenosti i neodređenosti te je njezino izdvajanje za potrebe jezične analize izazovan zadatak. Također, jezici bez sustava članova, poput hrvatskoga, koriste raznorodne strategije za izražavanje neodređenosti i određenosti, kao i njima srodne kategorije specifičnosti. U radu se tako raspravlja s kojim se pokazateljima određenosti i neodređenosti u hrvatskome može povezati i ostvarivanje specifičnosti. Tri kategorije za izražavanje određenosti i neodređenosti koje se obrađuju jesu: a) kratki i dugi oblici pridjeva, b) broj *jedan* i c) pokazne i neodređene zamjenice. Na primjerima iz Hrvatskoga nacionalnog korpusa te Hrvatske jezične riznice pokazuje se nestanak opreke neodređenosti i određenosti u kratkim i dugim oblicima pridjeva te u skladu s time i nedostatnost ove tradicionalne gramatičke podjele za opis kategorije specifičnosti. Posebna se pažnja posvećuje funkcijama broja *jedan* i procesu gramatikalizacije kojime *jedan* u određenim kontekstima gubi svoju funkciju broja te preuzima funkciju neodređenoga specifičnog člana. Nadalje, kao dodatni pokazatelj specifičnosti u hrvatskome ističe se složena konstrukcija *taj+neki,* odnosno konstrukcija pokazne zamjenice i neodređene zamjenice u poziciji modifikatora imenske fraze. Na temelju korpusne analize primjera i sredstava kojima se izražava specifičnost u hrvatskome ustanovljena je klasifikacija jezičnih sredstava za izražavanje specifičnosti koja se sastoji od kategorija a) kontekstualne specifičnosti, b) poluleksikalizirane specifičnosti i c) konstrukcijske specifičnosti. Kontekstualna specifičnost se definira kao pragmatički ostvarena kategorija, dok se poluleksikalizirana specifičnost ostvaruje gramatikalizacijom broja *jedan.* Konstrukcijska je specifičnost izražena slaganjem zamjenica *taj+neki.*

43