



## **Planning Against Change**

### **Serbian and Croatian reactions to contact-induced linguistic innovation**

Madsen, Martin Schou

*Publication date:*  
2017

*Document version*  
Other version

*Document license:*  
[CC BY-NC-ND](#)

*Citation for published version (APA):*  
Madsen, M. S. (2017). *Planning Against Change: Serbian and Croatian reactions to contact-induced linguistic innovation*. Det Humanistiske Fakultet, Københavns Universitet.



## PhD Thesis

Martin Schou Madsen

## Planning Against Change

Serbian and Croatian reactions to contact-induced linguistic innovations

## Contents

1. Introduction .....	1
Structure and methodology .....	6
1.1 The relatively short history of Serbo-Croatian .....	9
1.1.1 Unification and diversification .....	10
1.1.2 The first steps towards a common norm.....	13
1.1.3 Dictated commonality/unity (The interwar period) .....	19
1.1.4 The Second World War – Diversification.....	20
1.1.5 Communist era – Unification, diversification, separation.....	23
1.1.6 Partition - Independency .....	30
1.2 Differences.....	34
1.2.1 Scripts, pronunciation and spelling rules.....	36
1.2.2 Lexemes, semiwords and affixes.....	40
1.2.3 Morphosyntactic differences .....	42
1.3 The linguistic phenomena under investigation .....	44
1.3.1 Nouns.....	44
1.3.2 Nominal compounds.....	44
1.3.3 Semiwords and affixes .....	44
1.3.4 Indefinite articles.....	45
1.4 The empirical data.....	45
2. Language Planning and the Standard Language .....	57
2.1 Standard language and norm .....	59
2.1.1 System, norm, speech.....	65
2.2 Identity and communication .....	69
2.2.1 Why hamper communication? .....	72
2.2.2 One language, one nation.....	74
2.2.3 Purism .....	76
2.3 Language status and language function .....	85
2.3.1 Status, Corpus, Prestige and Acquisition Planning .....	86

2.4 Planning of Serbian, Croatian and Serbo-Croatian .....	90
2.4.1 Selection .....	94
2.4.2 Description and Prescription .....	96
2.4.3 Elaboration .....	100
2.4.4 Acceptance and Implementation .....	104
2.4.5 Expansion and Cultivation.....	108
2.4.6 Evaluation and Reconstruction .....	109
2.5 Language planning and Contact linguistics .....	111
3. Language Change and Contact .....	113
3.1 Perspectives on change .....	113
3.2 Language change .....	114
3.2.1 Grammatical change and lexical change .....	118
3.2.2 Types of innovation.....	128
3.3 Contact-induced language change.....	132
3.3.1 Types of contact-induced innovations.....	133
3.3.2 Hierarchies or What is most replicable? .....	142
3.3.3 The assumption of contact .....	143
4. Investigated Linguistic Innovations .....	146
4.1 Recent research in lexical divergence between Serbian and Croatian concerning replicated matter.....	146
4.2 Adaptation of replicated matter.....	154
4.2.1 Normative attitudes towards replicated matter.....	159
4.2.2 Substitution of replicated matter.....	162
4.3 Nouns.....	167
4.3.1 Empirical data .....	171
4.4 Nominal compounds.....	180
4.4.1 The problem of categorisation.....	182
4.4.2 Contact-induced attributive, subordinate and coordinate compounds in Serbian and Croatian .....	193
4.4.3 Empirical data .....	204
4.5 Semiwords and affixes .....	220

4.5.1 Formal and semantic features of semiwords and affixes .....	224
4.5.2 Semiwords .....	226
4.5.3 Derivational affixes .....	234
4.5.4 Empirical data .....	240
4.6 Indefinite articles .....	252
4.6.1 Indefiniteness in articles .....	259
4.6.2 Research of article-like usage of <i>jedan</i> and <i>neki</i> .....	263
4.6.3 Empirical data .....	266
5. Recapitulation and Conclusions .....	279
5.1 Language planners and language users .....	279
5.2 From lexical to grammatical .....	281
5.3 Symbolic and communicative value .....	285
Appendices .....	289
Appendix A .....	289
Appendix B .....	302
Appendix C .....	324
Appendix D .....	332
1. Querying nouns .....	333
2. Querying nominal compounds .....	334
3. Querying semiwords and affixes .....	338
4. Querying <i>jedan</i> and <i>neki</i> .....	340
Bibliography .....	342
Resumé .....	360
Summary .....	361

## List of Figures

Figure 1.1	Countries where Serbo-Croatian was the majority language.....	9
Figure 2.1	Standard language according to Silić.....	67
Figure 2.2	Language on five levels of abstraction.....	68
Figure 2.3	Radovanović's ten phases of language planning.....	91
Figure 2.4	Haarmann's ideal typology of language planning.....	91
Figure 2.5	Hornberger's integrative framework of language planning typologies.....	93
Figure 3.1	From nominal phrase to nominal compound.....	139
Figure 3.2	From numeral to indefinite article.....	140
Figure 4.1	Duckworth's categorisation of matter substitution.....	165
Figure 4.2	Six possible compound structures.....	188
Figure 4.3	A continuum from lexeme to affix.....	230
Figure 4.4	Indefinite properties of <i>jedan</i> and <i>neki</i> in NPs.....	267

## **Acknowledgements**

First of all I wish to thank my friends, family and colleagues for providing me with moral and practical support during the years of my PhD-education.

I wish to acknowledge the immensely motivating impact my first supervisor, Jens Nørgård-Sørensen (†2015), had on me.

I also wish to thank my second supervisor Frans Gregersen for stepping in and continuously helping me through the challenges and pitfalls of the writing process.

Further, I wish to extend my gratitude to University of Copenhagen, Faculty of Humanities for the financial support that enabled the writing of this thesis.

Finally, I want to thank my wife and pillar in life, Helle who has supported me all the way.

## 1. Introduction

The focus of this dissertation is on two language communities' reactions to the same global challenge, viz. the Anglification of languages as part of the globalisation process<sup>1</sup>. The two language communities are the Serbian and the Croatian ones. During a long process, starting in the 19th century, a common standard Serbo-Croatian was established as the written standard for both Serbs and Croats (with certain modifications introduced below). Thus it became common to speak of Serbo-Croatian as one language. The official tie between Croatian and Serbian was severed during the civil war which followed upon the dissolution of what was once Yugoslavia. Serbo-Croatian was not only deprived of its status, it also ceased to exist as an official language.

The present is particularly interesting because in general, languages “seem to change at certain historical stages more abruptly than otherwise.” (Moser 2013: 104) The late 80s, the 90s and 00s represent such a historical stage, during which Croatian and Serbian underwent changes motivated by not only the hostility and wars among Croats and Serbs but also by the on-going globalisation of also this part of Europe.

The empirical focus of this dissertation is on language change in Serbian and Croatian due to the influence from so-called prestigious international or global languages. I have chosen to narrow the scope of linguistic innovations to contact-induced innovations as these innovations not only mirror the fact that globalisation is a very discernible influence, but also because the reactions of both the linguistic authorities and the actual speakers and writers of Serbian and Croatian are generally known to have different views and practices in this respect.

Upon having discussed and defined the intricate workings of language planning of Serbian, Croatian and Serbo-Croatian and language change per se, I will therefore investigate specific linguistic innovations in texts written in Serbian and Croatian in the years immediately before, under and after the tumultuous events which led to the break-up of the state of Yugoslavia in the early 1990s, as documented in two large elec-

---

<sup>1</sup> In this study, the term *globalisation* is used in a sense corresponding to Scholte's definition of a) internationalisation, i.e. increased interaction and interdependence between people in different countries, b) universalisation, i.e. the spread of something (or someone) to all corners of the world and c) deterritorialisation, i.e. the spread of supraterritorial transborder-connections and the weakened geographical territory of a country, or in this case, of the national domain of a language (Scholte 1999).



tronic text corpora, The Croatian National Corpus and The Corpus of Contemporary Serbian. In order to shed light on these issues, I ask the question:

Do the linguistic effects of globalisation enhance or diminish the differences between Serbian and Croatian?

This is my main and overarching research question to which I will propose an answer by pursuing the answers to the following questions:

What sort of impact does the language planning and language policy of linguistic authorities concerning contact-induced linguistic changes have on language usage? To what degree are the efforts of planners and policy makers to promote or demote usage of such innovations successful? Which kind of innovation is mostly warned against and which kind is less noticed? Do these particular linguistic innovations bring Croatian and Serbian closer to or further away from each other? Do any of the linguistic innovations constitute a change of the languages themselves?

The case of Serbian and Croatian provides a unique situation in which to pursue the answers to my research questions. Croatian and Serbian have a specific internal relationship in as much as they, because of their similarities, have been (and are still by some) considered as one language – Serbo-Croatian, a result of deliberate planning and policy, whereas now they are officially considered as two separate languages, also a result of deliberate language planning and policy. Their linguistic similarity coupled with their political and official diversity gives us an opportunity to study the mechanisms leading to these opposites. As a part of the globalisation process, linguistic Anglification, the ever growing influence and impact of the lingua franca of today, English seems to be pulling Serbian and Croatian back together again, as the language users, the speakers of these languages, share the political reality of being small communities with a need and wish to communicate with the same foreign political and commercial powers. They are thus influenced by the same cultural forces of the English-speaking world. So, similar languages, dissimilar language policies, existing under similar circumstances provide a prism through which I will show how contact-induced linguistic changes are received by both linguistic authorities and the language users themselves. In other words, the diversification of Croatian and Serbian is countered by the uniting force of globalisation and I will show whether or not the reactions to globalisation, visible in Croatian and Serbian, act as a catalyst of diversification or unification.

It is commonly accepted that standard Croatian is more resistant to loanwords than standard Serbian. This resistance is, indeed, often mentioned in descriptions of the dissimilarities between Serbian and Croatian and therefore Croatian is considered more puristic than Serbian. As an experienced teacher of Bosnian/Croatian/Serbian, I can testify, that serious research into this particular subject is sorely missed. This is not so much a matter of differences between the normative standards but more an issue of possible differences in usage norms in the five functional styles of each language. Without rejecting the fact that loanwords are less readily accepted into Croatian, my *first hypothesis* is that within other contact-induced linguistic changes the differences between the two languages lie not in the speakers' readiness to accept the changes but rather in the degree of their adherence to the recommendations of the normative bodies of the two nationalities. Setting aside the very visible signs of foreign influence in the lexis, my *second hypothesis* is that linguistic innovations in Serbian and Croatian, induced by contact with dominant languages, diminish the dissimilarities between Croatian and Serbian, that is, the effects of globalisation on these two languages are more uniting, than they are dividing, despite different actors' efforts to diversify Croatian and Serbian by adapting to linguistic globalisation in different ways.

The pursuit of answers to my research question has led me to address the *general question* of how contact with other languages induce changes in language usage as well as to the more *specific question* of which contact-induced changes are more acceptable to linguistic authorities and language users in the two linguistic communities. As I am dealing with not only the dichotomy Serbian : Croatian, but also Standard language norm : Usage norm, the concept of *norm* is crucial to my approach. Therefore a discussion of this concept and a clear and applicable definition of it is also a vital part of this dissertation.

As I have investigated lexical as well as grammatical changes, my research will also contribute to the *discussion of language change* per se, more specifically:

When does a change in language usage constitute a change in the language system? Or: When is a language change merely a lexical borrowing? When may we conclude that an innovation is enhancing the word-formational possibilities? And when is an innovation changing the very morphosyntactic possibilities within a language?

By viewing the linguistic innovations in their societal context I aspire to follow the ideal put forward by Thomason (2003) that “both social and linguistic factors must in principle be considered in any full account of any linguistic change” (Thomason 2003: 688).

Throughout the thesis I draw on general frameworks and theories of language change<sup>2</sup>, language planning and policy, including questions of norm and standardisation<sup>3</sup>, on descriptions of the linguistic categories under investigation<sup>4</sup> as well as existing research within the field of contact linguistics pertaining to the influence of dominant languages on other languages in general and specifically; on Serbian, Croatian, Bosnian and Serbo-Croatians<sup>5</sup>.

### *Serbian, Croatian and Serbo-Croatian*

In spite of the current political definition of Croatian and Serbian as two distinct languages, they were, as mentioned above, until approx. 25 years ago, considered to be one language – Serbo-Croatian – and in the first half of the 20th century their speakers, Serbs and Croats, were even considered to be, and declared by their government to belong to, the same nationality. In order to navigate this, linguistically speaking, rather unusual situation, this thesis offers an introduction to the history of the linguistic relationship between Serbs and Croats. This introduction (section 1.1, p. 9ff) is based on research of differing descriptions of this history provided by a range of Croat and Serb linguists as well as linguists and other scholars with no apparent national ties to the region.

The relation between the two peoples and with them their language has had a significant influence on the development of standard Serbo-Croatian, Serbian and Croatian. The standard Croatian and standard Serbian are linked to each other both with regards to the linguistic structure and the socio-political circumstances. Some will say that this relation has helped along the development of a true polyfunctional standard language;

---

<sup>2</sup> Henning Andersen, Bernd Heine, Tania Kuteva, Laurel Brinton, Elizabeth Traugott, Paul Hopper, Jens Nørgård-Sørensen, Lars Heltoft, Lene Schøsler, Alice Harris, Robert Murray, Sarah Thomason

<sup>3</sup> Einar Haugen, Heinz Kloss, Eugenio Coseriu, Nancy Hornberger, Milorad Radovanović, George Thomas, John E. Joseph, John Edwards, Jozef Neustupný, Bohuslav Havránek, Harald Haarmann, Robert Kaplan, Richard Baldauf, Michael Clyne, Josip Silić, etc.

<sup>4</sup> *Among others*: Laurie Bauer, Rochelle Lieber, Sergio Scalise, Martin Haspelmath, Christopher Lyons, Ivan Klajn, Stjepan Babić, Predrag Piper, Eugenija Barić, Milka Ivić

<sup>5</sup> *Among others*: Goebel, Raymond Hickey Manfred Görlach, Fredric Field, Yaron Matras, Branislav Brborić, Ranko Bugarski, Branko Tošović, Barbara Štebih-Golub, Ivo Pranjković, Radoslav Katičić, Anita Peti-Stantić, Keith Langston, Rudolf Filipović

others will say that the development of the languages has been hampered by the relation. Today, besides having a shared past, the two standard languages also share a common base-dialect. The base-dialect of both standard Serbian and Croatian is the macro-dialect of Neoštokavian. In the mid-19<sup>th</sup> century Neoštokavian was selected by both Serbian and Croatian scholars as a basis for a standard language and this dialect was spoken by a large amount of Croats and Serbs living in what are today the independent states of Serbia, Croatia, as well as Bosnia-Herzegovina, Montenegro and, to a lesser extent, Kosovo.

Serbian and Croatian (along with Montenegrin and Bosnian) belong to the group of South Slavic languages which also include Slovenian, Macedonian and Bulgarian. The South Slavic languages can be subdivided into the western (Slovenian, Serbian/Croatian/Bosnian/Montenegrin) and eastern (Macedonian, Bulgarian) branches. The South Slavic area, wherein Serbian and Croatian are spoken, stretches from the Slovenian area in the Northwest to the Macedonian and Bulgarian area in the Southeast (Brozović and Ivić 1988: 2). Within the Serbian and Croatian area several dialects are spoken, the four major ones being: Kajkavian, Čakavian, Štokavian named as such with reference to the pronunciation of the lexeme *what* (kaj, ča and što) and Torlakian. Štokavian<sup>6</sup> is furthermore subdivided into the Neoštokavian and the Non-neoštokavian *and* into the versions: Ekavian, Ijekavian and Ikavian (named according to the different development in pronunciation of the Common Slavic vowel *jat* (Latin script: ě) (Cyrillic script: ѣ) into three major reflexes: e, (i)je and i).

The common linguistic foundation of modern standard Croatian and Serbian is today, as mentioned, Neoštokavian. So, out of four possible macro-dialects, Croats and Serbs selected the same: Štokavian. Out of several different variants of Štokavian, Serbs and Croats selected the same: Neoštokavian. Only when it came to the phonetic tripartite micro-dialectal subdivision of Neoštokavian the choices of Serbian and Croatian language planners differ in as much as the Serbs ended up standardizing not only one but two (of the three) microdialectal subdivisions: Ekavian and Ijekavian, whereas the Croats ended up standardizing only one of the them: Ijekavian. Modern normative standard Croatian is thus pronounced and spelled according to the Ijekavian version whereas modern

---

<sup>6</sup> Spoken by approx. 75% of all speakers of Croatian, Serbian, Bosnian and Montenegrin (Ivić 1994: 85)

normative standard Serbian is pronounced and spelled either according to the Ijekavian or the Ekavian version. The latter was also the case for standard Serbo-Croatian.

### **Structure and methodology**

I will support my hypotheses that non-salient linguistic innovations are more easily accepted and that globalisation unites rather than divides by investigating five types of presumed contact-induced linguistic innovations in Croatian and Serbian. I will do this by analysing findings in two electronic text corpora, one Croatian and one Serbian, which I consider representative of the (written) language usage norm among Croats and Serbs in 90s and 00s and by comparing these findings to the normative standards of the two standard languages.

The specific types of linguistic innovations that I have investigated comprise linguistic matter (lexemes, semiwords and affixes), word-formational patterns (in derivation and composition) and grammatical patterns (indefinite marking). These types of language change cover at the one end very salient changes which however do not change the structural make-up of the linguistic system of Serbian and Croatian (lexical innovations) and at the other end a possible fundamental change of their structural make-up (grammatical change).

Croatian and Serbian are, as mentioned, linguistically very similar but contact-induced linguistic innovations are very differently treated. In other words, the dissimilarity between Croatian and Serbian, that I am concerned with, is the dissimilarity in which contact-induced changes are deemed acceptable or not. The question of agency naturally arises and I will therefore also consider who it is that *deems* and *treats* the innovations in this way.

As a contrast for the analysis of empirical data, I will give an outline of the societal circumstances in which Serbian and Croatian have existed in the past one and a half century (1.1) as well as a description of the often mentioned, but in absolute numbers very few, linguistic differences between the two languages (1.2). In 1.3, you will find examples of the innovations under investigation followed by an introduction to sources of empirical data in 1.4.

Chapter 2 is devoted to a discussion of *language planning*, a phenomenon studied in *sociolinguistics*, which is the branch of linguistic research which focuses on the speech community. A speech community may be as small as a village or a club or as large as a nation or even a group of nations (Southerland and Katamba 1997: 540). In this chapter

on language planning I will, thus, by way of different frameworks, illustrate how the national standard languages Serbian and Croatian were, have been and indeed still are formed through this process. I will take the reader through the phases which a language variety usually goes in order to be promoted to the function and status of a *standard language*. Different views on language planning will be presented and the crucial part of the planners, the *agents* of planning will be illustrated and exemplified by outlining the Croatian and Serbian language planning and policy. In other words: How did the Serbian and Croatian standard languages come about? And who decides what is proper Serbian and Croatian? When discussing language planning the question of what the outcome of this planning should be, the question of *norm* and *standard language* naturally arises. To answer this question I will discuss Eugenio Coseriu's view of linguistic norm and the Croatian sociolinguist Josip Silić's elaboration of it (Coseriu 1974; Silić 2006).

Chapter 3 is devoted to the vast field of language change, but I focus on language changes induced by contact with other languages. Language change is traditionally studied in the field of historical linguistics where changes in linguistic structures, which have happened through a period of several hundred years, are studied. The motivation for the changes (the human factor) is usually not the focus of language change studies. By focusing on contact with specific languages, and on a period of great societal changes I strive to apply the theoretical findings from scholars in historical linguistics. I will therefore draw upon a general theory of language change by Henning Andersen (2001, 2008), and enlist the aid of among others Bernd Heine and Tania Kuteva (2005, 2006, 2010), Hopper and Traugott (2003) as well as Brinton and Traugott (2005) in discussing the phenomena of grammatical and lexical change (3.2). Language change is, however, also viewed as on-going process (a gradual change), so subsequently I will add the insights into gradual contact-induced language change (3.3) offered by Sarah Thomason (2001, 2003), Bernd Heine and Tania Kuteva (2005, 2006, 2010) Henrik Gottlieb (2006) and Yaron Matras (2009, 2010). Finally, I will touch upon the problems in identifying a change as having been induced by contact with another language and not simply a change which has happened motivated by other, language-internal factors.

In principle, every time a linguistic innovation is made, the inventory of the language is changed. The replication of lexical matter (loanwords) does not alter the structure of the language as such but it cer-

tainly enriches the lexicon. When word-formational patterns are replicated, structural changes within word-formation in the receiving language may occur and if a grammatical category is replicated, the structure of the receiving language is likely to change. A replication of some linguistic matter (lexical or non-lexical) may in time instigate changes in the linguistic pattern (the structure) of the receiving language which may lead to a grammatical change.

Chapters 2 and 3 represent the theoretical frame in which the specific linguistic phenomena under investigation in chapter 4 are to be viewed. The reactions of language planners and the language policies towards so-called contact-induced language changes and not least the impact of language planning measures aiming at diminishing the influence of contact with other languages circumscribes scientific fields where theories of language change, contact linguistics and sociolinguistics meet. The historical overview and the explanatory frameworks of language planning introduced will serve as a necessary context in which specific language changes, linguistically explained by means of universal theories and models of language change, observed in Serbian and Croatian modern language usage, are analysed.

In chapter 4 I will describe, problematize and define the concrete phenomena under investigation, which are: replicated *nouns* and their suggested substitutes, replicated *compound structures*, replicated *semi-words* and *affixes* and replicated *indefinite marking* (for examples, see section 1.3 below. Subsequently, I will present the methods used and the results gained from the empirical analysis of data<sup>7</sup> found in the Croatian and Serbian 100m+ national electronic text corpora. I will relate my findings to existing research as well as to the preceding chapters of the thesis.

In chapter 5 I will conclude the thesis by summing up and relating my findings in chapter 4 to the language planning and policy efforts of Croatia and Serbia and present my conclusions on these specific contact-induced language changes in Serbian and Croatian. I will present my findings concerning the impact of the efforts of language authorities on language use specifically among Serb and Croat language users and present some general thoughts on the issue. I will finally summarise my answers to my other research questions concerning the uniting effect of globalisation as well as the question of whether any more profound structural change is going on in Serbian and Croatian.

---

<sup>7</sup> Appendices A, B and C contain lists of the empirical evidence of compound structures, semiwords and affixes.

## 1.1 The relatively short history of Serbo-Croatian

Serbo-Croatian<sup>8</sup> was the official language of four out of six republics within the Socialist Yugoslav Federation: The Socialist Republic of Croatia, The Socialist Republic of Bosnia-Herzegovina, The Socialist Republic of Montenegro and The Socialist Republic of Serbia. Today, in lieu of these four republics we find four sovereign states<sup>9</sup>. The national language of present day Croatia is Croatian, of Serbia Serbian, of Montenegro Montenegrin and of Bosnia-Herzegovina, Bosnian, Serbian and Croatian.

It could be claimed that these new names are but a result of a classification made by speakers which differs from that of professional linguists. This type of classification is seen as an emic classification because of the languages' symbolic and social status in the minds of their speakers (Goebel et al. 1997a: 1975)(Goebel et al. 1997b: 2: 1975). The Serbian and Croatian official languages are the two of these four official languages which are most distinct from each other and they may be viewed as representatives of two extremes within the same dialect, i.e. Neoštokavian, which has Croatian at its most western point and Serbian at its most eastern point.(Vasyl'eva 2009: 165). This correlation mirrors the geographical areas of present-day Croatia and Serbia, i.e. Croatia covers the western and northern part of this dialect continuum while Serbia covers the eastern and southern part, cf. Figure 1.1 below.

*Figure 1.1 - Countries where Serbo-Croatian was the majority language*



10

---

<sup>8</sup> *srpskohrvatski* also called *hrvatskosrpski* (Croato-Serbian), *hrvatski ili srpski* (Croatian or Serbian)

<sup>9</sup> A 5<sup>th</sup> state, Kosovo (formerly a province in Serbia) has been recognized as an independent state by the majority of the international community.

<sup>10</sup> <http://kingofwallpapers.com/yugoslavia.html> (Accessed Oct 15, 2016)



Croatian and Serbian also each has a literary tradition which goes back to medieval times, i.e. long before the term Serbo-Croatian was coined, but also long before any kind of normative standardisation of the language(s) had been carried through. This means that text produced in Croatian and Serbian before the standardisation, today are labelled Croatian or Serbian not so much because of their linguistic relation to either Serbian or Croatian modern standard language but in relation to who wrote the text which in turn means that the origin of the author or text dictates the 'nationality' of the language. Quite often, the geographical origin of the author is not regarded as a sufficient parameter. The ethnicity (nationality) is also seen as a decisive factor. This means that the modern nations/nationalities often fight over the ownership of older canonized works or writers. This goes for the famed 17<sup>th</sup> century literature of Dubrovnik (a part of present day Croatia) but also regarding modern (20<sup>th</sup> century) writers such as Nobel prize winner Ivo Andrić<sup>11</sup> who grew up in Bosnia with Croat parents, but moved to Serbia (Belgrade) and wrote his main works there in the Eastern (Serbian)<sup>12</sup> variant of the Serbo-Croatian language.

The ideas of nationality and language are, as we have just seen, much intermingled. In the following subsections 1.1.1-1.1.5 I will provide an overview of the history of Serbo-Croatian from the first steps towards a common norm, in the mid-19<sup>th</sup> century to the official disintegration Serbo-Croatian at the end of the 20<sup>th</sup> century. It is a relatively short overview which is given here to provide the readership with an idea of the activities and circumstances that initially led to the integration and subsequent disintegration of the Serbian and Croatian standard languages.

### **1.1.1 Unification and diversification**

The main ideas which permeate the history of Serbo-Croatian, glottopolitically speaking, are the two opposing ideas of unification and diversification, respectively. In the beginning of the 19<sup>th</sup> century, as national awakening was flowing through Europe, Serbs and Croats did not share much. Present day central and southern Serbia was gradually becoming more and more independent of the Ottoman Empire while different areas of modern day Croatia along with modern day Northern Serbia was ruled

---

<sup>11</sup> The Nobel Prize in Literature 1961 was awarded to Ivo Andrić *"for the epic force with which he has traced themes and depicted human destinies drawn from the history of his country"*. ('The Nobel Prize in Literature 1961' 2016)

<sup>12</sup> The term variant, or Eastern variant was not put into use until 1965 (20-25 years later)

by different overlords mostly within the Habsburg (later Austro-Hungarian) Empire. It is not disputed even by firm believers in the unity of standard Serbo-Croatian that standard Croatian and standard Serbian until the 19<sup>th</sup> century consisted of two if not more distinct entities.<sup>13</sup>

Up until the beginning of the 19<sup>th</sup> century, literacy in Serbian was linked to the sacramental Church Slavonic of the Orthodox clergy and to the *slavjanoserbski* - a never standardized language of educated Serbs in Vojvodina, then Hungary, since 1918 part of Serbia and with a Serb population since the 16<sup>th</sup> century. Dositej Obradović (1739-1811) a protagonist of the Enlightenment believed that the language of the people (not the Church) would and should play an instrumental part in forming national identity (Sundhaussen 2009: 97). He thus wrote and translated literature into a language closer to the spoken language of the people than any other writer before him. This, consequently, made him unpopular within the Serbian Orthodox Church and it has been said that his books were bought by Serbian monks only to be burned (Jacobsen 1989: 29). Nevertheless, he came to be the founder of the forerunner of the University of Belgrade and in the year of his death he also became the first Minister of Education in the small Serbian de facto independent state (1804-1813).

At the same time, Obradović was admired by leading Croat protagonists of Enlightenment and popularization of the written language, i.e. Ljudevit Gaj and Stanko Vraz of the *Illyrian Movement*<sup>14</sup> (Jacobsen 1989: 29). Ljudevit Gaj is still recognised as one of the chief figures in the national awakening among Croats, the *hrvatski nacionalni preporod* (Croatian National Revival). The idea of promoting the vernacular to the status of literary (standard) language became central both amongst leading Croats and Serbs.

In this process of promoting the status of one particular version of the language, nationality came to be affiliated to one's vernacular language instead of one's religion. Instead of defining one's nationality on religious grounds (i.e. "I am a Serb, because I belong to the Serbian orthodox Christian community"), nationality was defined according to

---

<sup>13</sup> Pavle Ivić puts this point across by saying "... zajedničkog književnog jezika Hrvata i Srba nije bilo. Takvo je stanje napušteno tek u XIX veku, ..." (*there was no common literary language of Croats and Serbs. This state of affairs was not abandoned until the 19<sup>th</sup> century*) (Ivić 1990: 312)

<sup>14</sup> Centred around the first daily newspaper to be published in Croatian in January 1835 *Novine Horvatzke* (Croatian News) and its literary addition *Danicza Horvatzka, Slavonzka i Dalmatinzka*.

one's language (i.e. "I am a Serb, because I speak in the Serbian language."). This dichotomy has been termed linguistic vs. confessional nationality. (Sundhaussen 2009: 98, 103). In this respect the Croatia-based Illyrian movement faced a great challenge. The Croats, even though they all belonged to the Catholic faith, did not have one vernacular but three. Croats spoke (and wrote) in the three distinct western south Slavic dialects<sup>15</sup>: *Štokavian*, *Kajkavian* and *Čakavian*, cf. p. 5.

Čakavian was the main dialect in which literature was written in the Renaissance and Baroque era (16<sup>th</sup> and 17<sup>th</sup> century) when poets in the region of Dalmatia (on the Adriatic coast) took up writing in their mother tongue (as opposed to Latin or Italian). Kajkavian is the dialect spoken in Zagreb and the surrounding Zagorje and as Zagreb became a cultural centre of Croats in the 17<sup>th</sup> century, Kajkavian correspondingly became the most prestigious vernacular. Neoštokavian had by the 19<sup>th</sup> century replaced written Čakavian in most places (Brozović and Ivić 1988: 100) and was also in use among literary Croats in the city state of Dubrovnik (on the southern Dalmatian coast) and in the rural Slavonia (the easternmost part of Croatia).

Another challenge for the Illyrian movement, active from 1835, was what today seem to be two incompatible objectives. On the one hand, it was a movement which strove towards national revival of the Croats but on the other hand it promoted the idea that all Southern Slavs should come together (Mønnesland 2002: 20). Faced with the dilemma of the three dialects, Ljudevit Gaj and with him The Illyrian movement opted for the Neoštokavian dialect (the macro-dialect of the idioms of Dubrovnik, Slavonia, Bosnia-Herzegovina, Serbia and Montenegro) and thus cleared the road so that all Southern Slav dialects would inevitably "...merge into one modern literary language" (Despalatović 1975: 82). Ljudevit Gaj might have cleared the road but, as we will see, it is questionable whether the dialects merged into one modern language.

The process of choosing one base-dialect upon which to build a literary (standard) language is the initial stage of what, in the language planning literature, is called *selection of a norm* (Haugen 1966: 18). And this was exactly what was done in the Illyrian movement when Neoštokavian was selected. In the relevant successor states of Montenegro, Croatia, Serbia and Bosnia-Herzegovina today, I would claim, a unification of languages

---

<sup>15</sup> Not dialects *of* a language, but dialects as in not (yet or fully) standardised language varieties

with different national designations is viewed as more or less synonymous with the merging of the nationalities themselves as indeed it was in the 19<sup>th</sup> century when the first steps towards linguistic unification were taken. It may then seem quite logical, that today's countries with their newly found political independence, autonomy and sovereignty do not wish for this autonomy to be symbolically watered down by admitting linguistic commonality with other nationalities and states. The opposite, a deliberate diversification of how *we* and *the other* speak and write is rather to be expected. It is evident that the existence and promotion of a national language go hand in hand with the promotion and (re)building or revival of national identity. This has also been observed elsewhere in the world. Kulick (1992) quoted in Thomason (2001: 84) tells us that "New Guinean communities have purposely fostered linguistic diversity because they have seen language as a highly salient marker of group identity." Or, as Leach (1964) puts it in his study of Burmese tribes and the political systems to which they adhere: "[T]o speak the same language as one's neighbours expresses solidarity with those neighbours, to speak a different language from one's neighbours expresses social distance or even hostility." (Leach 1964: 49)

### 1.1.2 The first steps towards a common norm

The first steps towards a united standardisation of Serbian and Croatian were made in the mid-19<sup>th</sup> century by a small group of idealistic language reformers and writers from the Serbian and Croatian cultural elites, when they met in 1850 and on March 28 signed the *Vienna Literary Agreement* 67 years before most of modern day Serbia and Croatia were to be part of one country (The Kingdom of Serbs, Croats and Slovenes (1918)).

*"We, the undersigned, aware that **one** people needs to have **one** literature, and in that connection with dismay witnessing how our literature is splintered, not only by alphabets, but still by orthographic rules as well, have convened these past days to discuss how we could agree and **unify** in our literature as much as is now possible."* [emphasis added] (Translated by Greenberg in Greenberg (2008: 183))<sup>16</sup>.

---

<sup>16</sup> Original: "Dolje potpisani znajući da jedan narod treba jednu književnost da ima i po tom sa žalosti gledajući kako nam je književnost raskomadana, ne samo po bukvi nego još i po jeziku i po pravopisu, sastajali smo se ovijeh dana da se razgovorimo kako bismo se, što se za sad više može, u književnosti složili i ujedinili."

These introductory lines of the Vienna Literary Agreement show that the leading idea was to unify nations and the unification of nations entailed a unification of languages.

The signatories had no real authority in their hinterland, but their agreement would set the ground for working towards a (common) goal of one standardised language. The significance of the Vienna Literary Agreement has, however, been disputed and Croatian linguist Radoslav Katičić claims that it was not until later in history, that this document was elevated to an important document and as such the starting point of the standardisation of a joint language of Serbs and Croats (Katičić 2008: 36).

In the Vienna Literary Agreement some fundamental language planning<sup>17</sup> steps were taken. A dialect (Neoštokavian) was selected and the signatories agreed “[...] that the language should be a unified standard for Serbs and Croats.” (Greenberg 2008: 24). The principles of the reform of Serbian were proposed by Vuk Stefanović Karadžić (1787-1864), a man whose ideas of a Serbian language of the people were to be the foundations of a linguistic reform of not only the Serbian but also the Croatian prescribed norm. The corpus of the common language of Serbs and Croats should be the language of the common (illiterate) people not the clerical or civic elite. Vuk Karadžić and his contemporary and subsequent followers wrote grammars and dictionaries on the basis of how common people talked and expressed themselves in folkloric poetry. This, in turn, meant that the description of the common language would be a description of the commoners’ language. The elevation of the commoner’s language led to the severing of the connection to the literature of the civic and cleric elite of the preceding centuries. In other words the language of the 16<sup>th</sup>, 17<sup>th</sup>, 18<sup>th</sup> and 19<sup>th</sup> century poets and writers was not part of the corpus, that was to become the basis of modern Serbo-Croatian. The description of the Neoštokavian macro-dialect spoken by the commoners is what in language planning terms is called *codification of form* which does not solely entail description of the selected idiom but also prescription of how it should be used, i.e. what is correct language use and what is not.

Despite opposition in both speech communities, the idea of a common, unified language gained ground in both Croatia and Serbia.

---

<sup>17</sup> The concept of ‘Language planning’ will be described and explained in chapter 2, below

### ***In Croatia***

In Croatia, philologists who supported the ideas of Vuk Karadžić worked towards linking the Croatian standard (literary) Neoštokavian language to Serbian. They concerned themselves with codification of the vocabulary, grammar, stylistics, orthography but also the cohesion with the Čakavian and Kajkavian Croatian dialects (Auburger 1999: 27).

Writers and poets who gathered around The Illyrian Movement and its periodical *Danica*<sup>18</sup> promoted the use of Štokavian and wrote their texts in Štokavian using a reformed orthography, later called *gajevica* named after Ljudevit Gaj. The codified standard Serbo-Croatian in Croatia on the basis of the Neoštokavian dialect was not, however, fully implemented until the 1890s by the orthographic dictionary: *Hrvatski Pravopis* by Ivan Broz (1892), the grammar by Tomo Maretić (1899) *Gramatika i stilistika hrvatskoga ili srpskoga književnog jezika* and Ivan Broz and Franjo Iveković' dictionary of Croatian (1901) *Rječnik hrvatskoga jezika*.<sup>19</sup>

The question of when the standardisation of Croatian began is disputed. The traditional view that the standardisation began in the 1830s as part of The Illyrian Movement was challenged by leading Croat linguists as early as the 1970s and standardisation of Croatian is now said to have begun in the mid-18<sup>th</sup> century when both Kajkavian and Štokavian were being standardised (Mićanović 2012: 98)<sup>20</sup>.

If standardisation is viewed as a process in which an idiom is elevated from regional to superregional status, i.e. an idiom becomes a means of communication cross-regionally; the standardisation process of Neoštokavian among Croats in the southern and eastern regions did begin in the mid-18<sup>th</sup> century as the Kajkavian standardisation was under way in the Northern and Western regions (Brozović 1978: 53).

In other words, when an idiom becomes the standard means of communication between regions, standardisation is already taking place. As this idiom gains ground and becomes polyfunctional, i.e. is used for

---

<sup>18</sup> The journal changed name several times and is today referred to simply as *Danica*. From 1836 the *Danica* was written in the štokavian base dialect. It was published 1835-1849 (Leto 2004: 164)

<sup>19</sup> This dictionary was heavily criticized among Croats for not using Croatian literature as empirical data, but rather the corpus of folklore that Vuk Karadžić had gathered as well as the poetry of the prince-bishop of Montenegro (1833-1851). Among Serbs it was heavily criticized for its title as it in their view was a dictionary of Serbian. (Pranjeković 1997:99)

<sup>20</sup> In the 70s and 80s Dalibor Brozović and Radoslav Katičić, in the 90s Milan Moguš, Dubravko Škiljan, Josip Vončina, Zlatko Vince

administrative, publicistic, literary and scientific communication, it is continually standardised.

Standardisation is, however, only a process that may or may not lead to the inauguration of a *standard language*, i.e. a standard language as the official means of communication in a given speech community.

Even after the selection of Neoštokavian had been broadly accepted, four different schools, three of them originating in urban centres (Rijeka, Zagreb and Zadar) and the fourth; the Croatian *Vukovci* ('Vukovites' - followers of Vuk Stefanović Karadžić) had opposing views on how to proceed to the next step in language planning - *the codification of form* (Ivić 1990: 125). In the end, the Vukovites won this particular glottopolitical battle, as the Croatian government<sup>21</sup> in 1889 commissioned Ivan Broz to write the above mentioned Croatian orthography *Hrvatski pravopis*, which was published in 1892 and which adhered to the principles of Vuk Stefanović Karadžić as did the also mentioned grammar by Tomo Maretić, published in 1899. (Ivić 1990: 125–26) and (Moguš 1995: 201).

The Croatian government introduced the 1892-orthography in schools in Slavonia and Croatia. In Dalmatia it was implemented soon after<sup>22</sup>. So, by 1914 only a few Croatian writers still adhered to the older morphophonological orthographies (Tomo Maretić 1914: 223). In 1913, Serbo-Croatian even gained the status of being the only official language in Croatia (not having to share this honour with German, Latin or Hungarian). Croatia was, at the time, a part of The Austro-Hungarian Empire but the decision that was passed in the Croatian parliament (Sabor) was never confirmed by the emperor due to the outbreak of the First World War (P. Ivić 1998: 278). A number of codification publications from the three opposing schools predating those of Broz and Maretić, representing views and normative rules concerning morphology, micro-dialect, spelling and even punctuation differing from those of the Vukovites as well as of each other also existed.<sup>23</sup>

The fact that the Vukovites in their codification works focused on the folklore sources that Vuk himself had gathered, thus leaving out a large mass of literary work from Croatian areas, both Štokavian and Non-

---

<sup>21</sup> 1868-1914 Croatia was under Hungarian domination, so the Croatian parliament "Sabor" answered to a *ban* (viceroy) appointed by the Hungarian Prime Minister. Especially in 1883-1903 under ban Khuen-Héderváry Croatia was subject to anti-Croatian politics. (Mønnesland 1999: 85)

<sup>22</sup> At the time, different parts of today's Croatia were ruled by different overlords

<sup>23</sup> For a brief overview of the controversies cf. Pranjković (2011) "Hrvatski jezik u 19. stoljeću"

štokavian, and ignoring the previous work made by members of the opposing schools created an imbalance between Serbian and Croatian heritage in the codification of the common language and the rectification of this imbalance has been of utmost importance in Croatian language planning ever since (Alexander 2006: 411).

While it is not disputed that the finalization of the standardisation of Neoštokavian happened in the late 1800s, it has been pointed out that earlier forms of literary (standard) Croatian and Serbian differ in many ways and that the acceptance of the Vuk Stefanović Karadžić's reform among Croats in 19<sup>th</sup> century was only possible because "Karadžić's model corresponded with the mainstream of standardization as it had come about among Croats." (Katičić 1997: 168). Brozović, (a proponent of viewing Serbian and Croatian as one standard language with two variants) puts it another way when he says that the selection of Štokavian was not done because Serbs and Croats wanted or needed a common standard language. It was the other way around. It was because Štokavian was selected by both Serbs and Croats that they ended up having the same standard language (Brozović 1970: 99). It is an issue of cause and effect. The need for commonality (a common standard language) was not the reason why Štokavian was selected, but its selection caused the commonality.

In other words, the histories of Croatian and Serbian standard languages are so different, that it was a(n un)lucky coincidence that the persons leading the language modernisation were of similar mind. (Greenberg 2008: 25)

### ***In Serbia***

In Serbia the situation stabilized somewhat earlier but only after heated discussions and polemics over whether to adopt the reformed version of literary (standard) Serbian, proposed by Vuk Karadžić as early as 1818. This reformed Serbian literary language was accepted as the official language of Serbia in 1868<sup>24</sup>.

Before 1868 it had supporters among some poets and writers. The first publication written according to Vuk's reformed Serbian language (not written by Vuk himself) came out in 1826 and prominent writers such as Branko Radičević (1824-1853) and Petar Petrović Njegoš (1813-

---

<sup>24</sup> Serbia obtained a sort of independency from the Ottoman Empire in 1815, but was officially a Principality until 1882. During this period Serbia's independency grew until the country's de facto independency from 1867 resulted in recognition in 1878, followed by status as Kingdom in 1882.



1851) wrote works in the reformed Serbian (P. Ivić 1998: 211 and 216) even though it was officially forbidden in Serbia proper to write according to Vuk's reformed orthography up until 1859 (Jacobsen 1989: 39).

Among Serbian philologists Vuk's work to this day enjoy a high and sometimes almost religious status: Vuk's last dictionary (1852) was to become the 'foundation of a new literary (i.e. standard) language and the bible of Serbian philologists.'<sup>25</sup> (P. Ivić 1998: 233). The implementation of the standard Serbian was done through schools via orthographic dictionaries: *Srpski pravopis za školsku upotrebu* by Čuturilo (1884) and *Srpski pravopis za srednje škole* by Petrović (1914) and grammars such as *Srpska gramatika* by Novaković (1894). (Brborić 2013: 173).

The differences in the spoken language in the cities of Zagreb and Belgrade were, in spite of being influenced by the common standard language, still very large. In Zagreb there was diglossia. The Štokavian standard language was used in formal settings whereas the original Kajkavian dialect was used in more informal settings (Ivić 1990: 130). As the Serbian intellectual elite all spoke Štokavian, a spoken version of the standard language soon emerged and at the turn of the century a particular Belgrade style of speaking had taken form (Ibid)

### ***The name of the unified language***

The name of the common language – the lingonym – never really stabilized. The name Serbo-Croatian was first used in 1824 by German philologist Jacob Grimm and diffused by the Slovene Jernej Kopitar (co-signer of the Vienna Literary Agreement) before the mid-19<sup>th</sup> century (P.-L. Thomas 2002: 319)<sup>26</sup> and it was used by grammarians from Zagreb in 1854 and 1859 (Lenček 1976: 49). But, as can be seen from the titles of the normative publications mentioned throughout this dissertation, before the Great War this lingonym had not been broadly accepted in neither Croatia (Maretić, Broz and Iveković) nor Serbia (Čuturilo, Petrović). Even the Croats in favour of the unification preferred to use the rather fuzzy *hrvatski ili srpski* (Croatian or Serbian) instead of the compound *srpskohrvatski* (Serbo-Croatian) or *hrvatskosrpski* (Croato-Serbian).<sup>27</sup>

---

<sup>25</sup> Original: „Он је постао темељ новог књижевног језика и библија српских филолога“ (P. Ivić 1998: 233)

<sup>26</sup> Other authors (such as Sundhaussen) claim that the name Serbo-Croat was not seen until 1867 (Sundhaussen 2009: 107)

<sup>27</sup> The hyphenated *srpsko-hrvatski* introduced by the Austro-Hungarian authorities in Bosnia-Herzegovina in 1907 (Šipka 2005: 410) never gained ground, probably partly because hyphenated compounds in Slavic languages may indicate opposition and/or

### 1.1.3 Dictated commonality/unity (The interwar period)

In 1918 a constitutional monarchy, The Kingdom of Serbs, Croats and Slovenes was established. The Head of state was the hitherto King Alexander of Serbia. From this day onward, the majority of Serbs and Croats lived in the same country which in 1929 became a totalitarian kingdom known as Yugoslavia under the absolute rule of King Alexander. Between 1918 and 1941 Croatian and Serbian was officially unified and called Serbo-Croatian and in 1929 codified in an orthographic manual for all primary, secondary and vocational schools *Pravopisno uputstvo za sve osnovne, srednje i stručne škole Kraljevine SHS* ((Samardžija 2008: 19), (M. Šipka 2005: 413))<sup>28</sup> as well as in the orthographic manual *Pravopis srpskohrvatskog književnog jezika* (1923) by Aleksandar Belić (Ivić 1990: 133).

These publications as well as others<sup>29</sup> are by modern Croat linguists regarded as attempts at forcing Serbian on Croats or at least as an underlining of Serbian authority and centralistic power in the years 1918-1941<sup>30</sup>. Indeed, people in Croatia did not necessarily follow these manuals, but kept writing according to older Croat spelling dictionaries<sup>31</sup> and there were Croat philologists/linguists who opposed the dictation (dictatorship) of Belgrade (Samardžija 2008: 19-20). So there was, during the interwar period as Kenneth Naylor remarks a “...lively debate [...] around the literary language...” in both Belgrade and Zagreb (Naylor 1980: 82).

This lively debate mirrors the fact that a common language under whichever name had not met with broad acceptance among Croats. This fact is supported by publications in the late thirties and early forties, which emphasised the differences of Serbian and Croatian, most notably *Razlike između hrvatskoga i srpskoga književnog jezika* (‘The Differences between the Croatian and the Serbian Literary Language’) (Guberina and Krstić 1977). Successful or not, the official language policy of the centralist government was that there existed only one language, namely Serbo-

---

division between its components cf. also Danish “de rød-hvide” [lit. The Red-whites], meaning the Red and Whites (two individual and different colours)

<sup>28</sup> There is a discrepancy between the referred works as according to Šipka (2005) the manual was printed only in Ekavian and in Cyrillic script which must be a mistake as Moguš (1995: 211) includes a picture of the front page of this manual, and it shows that the manual was also published in the Latin script.

<sup>29</sup> The language journal: *Naš jezik* (‘Our language’) was one of them (published from 1932)

<sup>30</sup> Cf. Auburger (1999) *Die kroatische Sprache und der Serbokroatismus*, p. 27-29

<sup>31</sup> codified by Broz (*Hrvatski pravopis* 1892) and republished by Boranić with minor alterations (*Pravopis hrvatskoga ili srpskoga jezika* 1921))

Croatian<sup>32</sup> and thus there was only one orthographic manual for all primary, secondary and vocational schools in the country, written in the Serbian variant as mentioned above.

The traditional, and at this point in time, primary script in use among Serbs was the Cyrillic one, whereas the Latin script was used by Croats, Slovenes, Jews and Moslems. The Cyrillic script was also present in school books in general and many teachers working in Croatia spoke the Serb variant. (M. Šipka 2005: 414–15). The whole period was for Croatian, in the words of Marko Samardžija, author of *Hrvatski jezik i pravopis od ujedinjenja do kraja Banovine Hrvatske (1918.-1941.)*<sup>33</sup> 'a time of constant trials and opposition to a politically supported, fiercely advocated and tenaciously implemented linguistic unitarism'<sup>34</sup> (Samardžija 2012: 7)

By the end of the interwar period the forces of unification had not won the battle against the forces of diversification; rather they had antagonized those whom they wanted to unite by attempting to unify them under existing (Serbian) rule and hereby fuelling the Croatian forces that did not see any benefits in the union with the Serbs, politically, glottopolitically, socially or otherwise.

#### 1.1.4 The Second World War – Diversification

##### *In Croatia*

Earlier years' partial non-acceptance of a unified language in Croatia culminated during The Second World War when the official language policy within the fascist Independent State of Croatia (*Nezavisna Država Hrvatska – NDH*) 1941-1945 aimed at "break[ing] completely with the functional development of Serbo-Croatian, [...]" (Moguš 1995: 214). In NDH, a German satellite state, the differences between Serbian and Croatian were emphasised and the most salient features of a written language, the vocabulary and the orthography were in focus. A number of laws on language were passed, condemning the use of words, which could be con-

---

<sup>32</sup> Actually, in the Constitution of 28th June, 1921 when the Kingdom of Serbs, Croats and Slovenes was established, the language Serbo-Croato-Slovenian was proclaimed as the official language, hereby mirroring the Herderian idea of one nation, one language. (Bugarski 2004: 26)

<sup>33</sup> *The Croatian Language and Orthography from the Unification till the end of Banovina Hrvatska (1918 – 1941)*

<sup>34</sup> Original: " ...neprestano bilo vrijeme kušnja i odupiranja politikom podupiranu, gorljivo zagovaranu i ustrajno provođenu jezičnom unitarizmu"

ceived as Serbian (Nuorluoto 2012: 104). First of all, the status of Croatian changed. Croatian was declared a language in its own right, not a dialect of another language or a language held in common with another nation. (Samardžija 2008: 47), citing the “Law on the Croatian Language”). Furthermore, the norms of spelling and writing were changed. Orthographic principles of morphophonological spelling from earlier, furthest from the Vukovian phonological spelling principles, were (re)introduced and the Cyrillic script was banned. A committee on Croatian language, its purity and orthography was formed. Its task was to purify the Croatian Language, i.e. remove the words which *were not in accordance with the spirit of the Croatian language* and *foreign* words. Within this committee there was much disagreement and its final recommendations after 1½ years of existence regarding the orthography was a compromise which included only a few concessions to the morphophonological spelling principles.<sup>35</sup> However, the authorities ignored the committee’s recommendations and decreed a shift from the well-established phonological spelling rules to the morphophonological ones and in the autumn of 1942 a manual, which was to be used in primary and secondary schools was published<sup>36</sup> followed by an orthographic manual in 1944<sup>37</sup>. (Samardžija 2008: 55). The authorities disseminated the new codification not only through manuals in schools, but also most ardently via the media, i.e. radio, newspapers and magazines.

Examples of the hitherto applied phonological (phonetic) and the introduced morphophonological (etymological, root-) spelling rules<sup>38</sup>:

		<u>phonological</u>	<u>morphophonological</u>	
<i>iz + pred</i>	>	<i>ispred</i>	<i>izpred</i>	- ‘in front of’
<i>svjedoč- + ba</i>	>	<i>svjedodžba</i>	<i>svjedočba</i>	- ‘diploma’
<i>otc + ev</i>	>	<i>očev</i>	<i>otčev</i>	- ‘father’s’

It is important to note that the shift from a Vukovian or at least Vuk-inspired phonological, albeit anti-Serbo-Croatian view on how Croatian should be codified to a morphophonological anti-Serbian view was not supported by the majority of Croat linguists. The decision to make the

---

<sup>35</sup> For a detailed analysis of the committee’s work, through their minutes, cf. Samardžija 2009: 48-55

<sup>36</sup> *Koriensko pisanje*, Hrvatski državni ured za jezik, Zagreb 1942

<sup>37</sup> *Hrvatski pravopis*, Hrvatski državni ured za jezik, Zagreb 1944

<sup>38</sup> The three words are pronounced: /‘ispred/, /‘svjɛdɔdʒba/ and /‘ɔtʃɛv/

shift had very clear and overwhelmingly political motives, and was reached by the then political elite in Croatia. The fact that the change was not only launched, but also implemented and to some extent accepted by Croats can be viewed as the natural extension of the events in the interwar period, when the power from Belgrade and its use of Serbian in state administration marginalized the Croatian lexis and lingonym, making a unification of Serbian and Croatian undesirable among the majority of literate Croats and Croatian philologists.

The Croats' apparently quick acceptance and implementation into the usage norm of strongly politically motivated changes of the prescribed norm of standard Croatian reoccurred in the 1990s, which we will return to below.

### ***In Serbia***

In Serbia, which was occupied by German forces, and ruled by the puppet government under general Milan Nedić, no specific language planning measures are mentioned in the works I have consulted. This period is often simply omitted while the Croat endeavours in separating Croatian from Serbian are frequently, if not always, mentioned.

Looking into the period of occupied Serbia 1941-1944 and the policies of the Nedić puppet government, provides an insight into the overall ideological principles and policies governed by these principles. In educational policy, formulated by the Ministry of Education in collaboration with the Academy of Sciences and other institutions in the *Srpski civilni plan* ('Serbian civic plan'), educators in primary and secondary schools were to instil the pupils with the proper nationalistic sense of what it meant to be a good Serb, e.g. in the study of Serbian language, literature and musical treasures (Petrov 2011: 97) but also with a "knowledge" about the communist agenda which, according to the policymakers, included casting aside all national traits, the crown, one's private property, family and religious beliefs. (Škodrić 2011: 155). Furthermore, the lingonym Serbo-Croatian was banned in as much as the official language of Serbia under Nedić was Serbian (Gröschel 2009: 19).

So it was not only in official Croatia but also in official Serbia that nationalistic projects were undertaken emphasising the importance of being a proper Croat or Serb, respectively. However, the nationalists, in whose interest diversity was, did not win the war. The communists, who emphasised the unity and uniformity at the cost of national treasures, did.

### 1.1.5 Communist era – Unification, diversification, separation

After the communist takeover, following the World War II, an official policy aimed at unification of Serbs and Croats was pursued once more. Initially, the Antifascist Council for the National Liberation of Yugoslavia (AVNOJ) recognised the existence of four languages, each representing one nation: Serbian, Croatian, Slovenian and Macedonian<sup>39</sup> (Moguš 1995: 216). However, Serbian and Croatian continued to be viewed and treated by the authorities as one language.

#### *Unification*

The idea of a united Serbo-Croatian was far from dead and thus an agreement, The Novi Sad Agreement, was signed in 1954, in which it is stated that the language of Serbs, Croats and Montenegrins is one language. In the agreement it is also acknowledged that this united (or unique<sup>40</sup>) language had developed around two centres (Zagreb and Belgrade) and that both parts of the lingonym (Serbian and Croatian) should always be mentioned in official use of the language. So, with regards to the lingonym, the supranational uniformity was pursued but at the same time the historical national diversity was recognised. The formal recognition of two centres for one language shows that the idea of pluricentric standard languages, as Kloss (1967) defines it was already present when the Novi Sad Agreement was drawn up. A pluricentric standard language is present in “those instances where we have two variants of the same standard. Serbo-Croatian is a case in point.” (Kloss 1967: 31). A pluricentric standard language is a language which develops in different centres. These centres can be located in different countries (e.g. Portuguese in Brazil and Portugal) but they can also be found within the same country “where political circumstances have brought about separated developments for two [or more] variants of one single language” (Ibid.).

The Novi Sad Agreement stated that the language needed a unified professional terminology, a unified orthography and a dictionary. Thus, classical standardisation measures pertaining to implementation of a standard, i.e. *codification of form* or *corpus planning* were taken.

---

<sup>39</sup> Macedonian had hitherto been viewed as either a dialect of Serbian or Bulgarian or a mixture of these two. Macedonian was declared the official language of the Macedonian Republic by the Anti-fascist Assembly for the National Liberation of Macedonia (ASNOM) in August 1944 (Koneski 1980: 62–63)

<sup>40</sup> *jedinstven* may mean both *unique* and *united* (Benson 1996: 183) (Bujas 2008: 488)

In the ensuing codification work, the compilation of a common orthographic manual (published in 1960) and a dictionary of Serbo-Croatian or Croato-Serbian, the individual entries were not marked as being either Croatian or Serbian, but simply as Serbo-Croatian/Croato-Serbian words. In line with the overall policy, emphasis was put on the unification of the diversities. This disregard of the origin of the different lexemes ultimately, led to the discontinuation of the work on a common dictionary in January 1971.<sup>41</sup>

### ***Diversification - One language, but two variants***

The lack of recognition of what was distinctively Croat or Serbian was, according to Moguš (1995), one of the reasons behind a stream of protests, which found its form in the *Declaration on the Name and Status of the Croatian Literary Language*<sup>42</sup> issued in March 1967 and co-signed by leading members of Croatian cultural and academic institutions (Moguš 1995: 220).

Already in 1965, in spite of the politically (communist) motivated work on strengthening the unity and unification of Serbian and Croatian formalised in the Novi Sad Agreement, linguists inspired by the sociolinguistic theories from Prague<sup>43</sup> succeeded in putting the diversity of how the common language of Serbo-Croatian was used if not on the political, then the scholarly agenda.

At the Fifth Congress of Yugoslav Slavists in Sarajevo in 1965 the two variants, the then coined “Eastern variant” and “Western variant”, of the Serbo-Croatian language were recognised (Katičić 2008: 32). Milka Ivić proposed and was supported by her colleagues in that instead of only prescribing what is proper Serbo-Croatian, on the basis of old and/or literary sources, the modern usage of the (standard) language should be described on the basis of the linguistic everyday practice, which is best illustrated in the language of the media (Ivić, M. 1965: 4).

The problem of discrepancy between what was the corpus of the normative works as opposed to the corpus of its elite speakers is a

---

<sup>41</sup> The work was being done in the two cultural institutions *Matica Srpska* and *Matica Hrvatska*, respectively. After lengthy discussion of the codification principles between the two institutions on, the editors of the *Matica Hrvatska*-edition handed in their letter of resignation in the autumn of 1969 and in January 1971 the cooperation of the two *Maticas* was discontinued (Jonke 1971: 379-383)

<sup>42</sup> Original: *Deklaracija o nazivu i položaju hrvatskog književnog jezika*

<sup>43</sup> Which concerns what constitutes a linguistic norm and a standard language being a reflection of modern usage as well as tradition and thus having a “flexible stability”. A discussion of this concept follows in section 2.1, below.

general problem of standard languages. Standard languages are according to Ammon (1989) traditionally derived from three sources 1) Dictionaries and grammars, 2) Texts and speech from the educated elite who represent the model speakers and writers in a society and 3) Governmental/semi-governmental bodies that continuously set up the normative rules for new occurrences in the language use, thus standardising the innovations (Ammon 1989: 88–89). The dictionaries and grammars are typically written by the members of the educated elite who work within the governmental and semi-governmental bodies, which dictate the norm of the standard language.<sup>44</sup>

In the case of Serbian-Croatian, the attempt to make common dictionaries and grammars succeeded as long as the model was the folkloristic, rural corpus that Vuk Karadžić had gathered in the 19<sup>th</sup> century. However, because the educated elites of Croatia and Serbia spoke and wrote distinctly differently not only from the original models, but also different from each other, the contemporary model speakers and writers used two distinct models (variants) of the standard language. This discrepancy led to disagreement between linguistic experts in the semi-governmental bodies of Matica Srpska and Matica Hrvatska on how to continuously standardise Serbo-Croatian resulting in the discontinuation of the common dictionary mentioned above and consequently to diversified and separate standardisation work.

The role of the media in spreading language usage was not only noted by Milka Ivić but also by the prominent Croatian co-signer of the Novi Sad Agreement, Ljudevit Jonke, when he in 1958 points out and criticizes that the Serbian Ekavian variant is almost the only variant in use in the common news broadcasts of Radio Zagreb and Radio Belgrade emitted to the whole of Yugoslavia. Thus a violation of the 4<sup>th</sup> point of the Novi Sad Agreement pertaining to the absolute equality of the two pronunciations Ijekavian and Ekavian was committed. (Jonke 1971: 288–90).

The euphemisms applied to denote what laymen would call Serbian and Croatian were many, and a need to differ between the two in linguistic research fields was definitely present. We find examples of this e.g. in Petar Skok's article on competing verbal suffixes from 1955 where he uses the newly introduced terminology of the 'two centres' Belgrade

---

<sup>44</sup> The role of the actors or agents who prescribe the norm as opposed to speakers who are to implement it will be discussed in chapter 2, below.



and Zagreb and calls what is specific for Croatian or Serbian 'Zagrebizmi' and 'Beogradizmi', resp. (Skok 1955: 43).

In the 1967 "Declaration on the Name and Status of the Croatian Literary Language" and in the debate of the two Maticas it is underlined that according to basic constitutional socialist principles every people/nation has the right to complete sovereignty and unlimited equality towards all other national communities. And it was this right that, according to the signatories of the Declaration, was being violated when the Croatian people was not allowed to name its own language. Regarding the status, the signatories called upon the government to consolidate the clear and unquestionable equality of four literary languages: Slovenian, Croatian, Serbian and Macedonian, because they saw a tendency toward a perceived need for not only a united but also a uniform administrative language. This language seemed to be equivalent to the distinctly Serbian variant of Serbo-Croatian. In other words Croatian or the Croatian variant of Serbo-Croatian was being marginalized and was not treated as an independent language on equal footing with Macedonian and Slovenian and the reason for this 'evil' lay in the commonality with Serbian (under the lingonym Serbo-Croatian) within which Croatian was losing ground to Serbian because a large majority of federal documents was written in Serbian (though frequently in the Latin script) which stemmed from the fact that the federal administration was situated in Belgrade. In harsher words, Croatian was in danger of being labelled a dialect of Serbo-Croatian and Serbian would thus become elevated to the "proper" or at least the more prestigious form of Serbo-Croatian. Brozović (1978) calls this mostly Serbian variant of Serbo-Croatian *the 'federal' language* ('savezni' jezik). (1978: 70).

One attempt at highlighting Croatian's distinctiveness and independence from Serbian was the Croatian orthographic manual and dictionary written by Stjepan Babić, Božidar Finka and Milan Moguš which was deemed unsuitable and separatist by the regime and as a result hereof not published as planned in Zagreb in 1971. It was subsequently published in London in 1972 and again in 1986 in a version without index, explanation of abbreviations and foreword and was only printed and published in its full format in 1990 in Zagreb (Babić et al.1990: 347). The 1967 'Declaration on the Name and Status of the Croatian Literary Language' as well as a Serbian reaction *Predlog za razmišljanje* ('A Proposal to Ponder') made by a group of Serbian writers were condemned by the regime as nationalistic propaganda and there were repercussions for the signatories of both documents. Thus,

attempts at diversification, i.e. the recognition and possible promotion of diversities, were blocked by the Yugoslav authorities.

### ***Separation – the lingonym in the Constitutions***

In spite of the harsh reactions from the regime, calls for the equality of Croatian or Croato-Serbian with Serbo-Croatian continued (for instance Moguš 1968, Pranjić 1968, Jonke 1969b) and seven years later, in 1974, when a new federal Constitution was passed, it became possible for each federal republic in Yugoslavia to determine which idiom should be used in its internal administration. The Socialist Republic of Croatia did not hesitate to choose the Croatian literary language (Alexander 2006: 415; Bugarski 2012: 227). Thus, the 1974 Constitution of the Socialist Republic of Croatia states that the official language in Croatia is ‘The Croatian Literary Language which is called Croatian or Serbian’ (Sofronić 1974: 315)<sup>45</sup>. In Serbia, however, Serbo-Croatian was still the preferred lingonym in internal administration (Brozović 1978: 72). Accordingly, the 1974 Constitution of the Socialist Republic of Serbia does not state anything about the official language of Serbia (Sofronić 1974: 618)<sup>46</sup>. It was in fact only in the constitution of Serbia that there was ‘no mention of any local trait of Serbo-Croatian used as the official language in the republic.’ (P. Ivić 1992: 102).

### ***Separation - From variants to languages (1974 to 1991)***

The majority of the literature on the history of Serbo-Croatian makes a temporal leap from the period from 1974 until 1991. In 1974 the naming of the official languages in the republics was turned into a matter for the individual republics – as opposed to being federal, and in 1991 the next major steps in status planning of Serbian and Croatian were made as a more or less natural extension of the declarations of independence from the Yugoslav federal state first by Slovenia and Croatia (June 25, 1991), then Macedonia (September 18, 1991) and finally Bosnia-Herzegovina

---

<sup>45</sup> In chapter 3 „Društveno-političko uređenje“ under the heading „Položaj radnih ljudi u društveno-političkom sistemu“ we find Article 138, which is quoted and translated in full in subsection 2.4.4 below (p.105)

<sup>46</sup> In chapter 2 „Društveno-politički sistem“ under the heading „1. Položaj radnih ljudi u društveno-političkom sistemu“ there is no mention of language. However, in the same chapter under the heading „7. Ostvarivanje ravnopravnosti naroda i narodnosti“ we find Article 146 which is quoted and translated in full in subsection 2.4.4 below (p.105)

(March 1, 1992)<sup>47</sup>. [Sočanac 2012] [Bugarski 2012] [Nuorluoto 2012] [Alexander 2006].

„Član 246. Ravnopravni su jezici naroda i narodnosti i njihova pisma na teritoriji Jugoslavije.“

(Art. 246. On Yugoslav territory the languages and alphabets of nations and minorities are equal)

There was no official language policy in Yugoslavia besides this declaratory statement in the 1974 Constitution, which said that all the languages of the nations and minorities are equal and the similar statements in the Constitutions of the republics as mentioned under the previously. However, the language spoken in Serbia and Croatia did not stop developing in the years between 1974 and 1991. On the contrary, the diversities were emphasised. Through the 70s and 80s both scholars and laymen took a great interest in the sociolinguistic problem of national variants. The issue was constantly on the agenda (Brozović 1990: 16) At the 12<sup>th</sup> Congress of the Federation of Slavistic Societies in Yugoslavia, held in Novi Sad in 1988, the division of standard Serbo-Croatian into variants was defined as one of if not *the* most important glottopolitical problem in Yugoslavia (Vasić 1990: 5). There was a constant struggle over the status of the variants vs. the common standard. A unified common standard was pursued by Belgrade e.g. in the wording of legal acts, which met with protests in Croatia (Brozović 1978: 13–23) and in Croatia the new status of Croatian as the administrative language of the republic was utilized in an attempt to remove the word Serbian from the 1974 Constitution of the Socialist Republic of Croatia (quoted above) which, unsurprisingly, was met with protest in Serbia (Ivić 1990: 323).

The protagonists of these two opposite trends were labelled *unitaristi* (Unitarians) and *nacionalisti* (nationalists) or in softer terms *integralisti* (integralists) and *separatisti* (separatists). The integralists (mostly Serbs from Belgrade) believed that the standardisation of the common Serbo-Croatian standard language had not yet been brought to

---

<sup>47</sup> In Kosovo, a referendum for independence was carried out in late September 1991, with a vast majority in favour of independence. However, Kosovo's declaration of independence of 1991 was recognized by only one UN member state: Albania. In 2008 Kosovo made another declaration of independence and is now recognized as an independent state by most countries in the world. Montenegro became an independent state in 2006 after hitherto having been part of the remnants of Yugoslavia under different names.

an end and that further diversification of the regional usages (not recognising them as variants) should be avoided. The separatists, mostly Croats, claimed that a unity between Croatian and Serbian usage had never existed, nor could this unity ever be accomplished. And as such Serbian and Croatian were different variants of a standard language and these variants were manifestations of the tradition as well as the modern needs of nations. (Baotić 2005: 437–38). The idea of variants and that they were functionally equal to languages was not new and had been uttered earlier in the Communist era, e.g. (Jonke 1965: 10) and (Brozović 1970: 35–36).

In the 1970s and 1980s in Croatia there were harsh reactions from the Communist Party as well as from individual communists to what they called *nationalism in language* which according to them was present in the school books for primary and secondary education. The focus was on vocabulary and the uncalled use of specifically Croatian words instead of the well-established federal vocabulary. (Pranjković 2006: 45–46)<sup>48</sup>. In the media, distinctly Croatian vocabulary, so-called *croatisms*, were suppressed and people who stood up against the suppression were professionally marginalized. In the army, there were even examples of prosecutions for using the Croatian *časnik* instead of the federal and Serbian *oficir* (Grčević 2002: 150; Oczkova 2010: 321). The federal vocabulary was generally characterised by internationalisms such as *sistem* ('system') and *muzika* ('music') and other distinctly Serbian words (Pranjković 2006: 47). The differences in vocabulary were then, and still are today, what is mostly emphasised as differential markers between Serbian and Croatian. In the last twenty years words like *sistem* and *muzika* have been very rarely used in Croatian media.<sup>49</sup>

The irreconcilable positions on the status of the variants of the Serbo-Croatian language mirrored the tendencies in political opinion about how Yugoslavia should be governed, i.e. whether it should have a strong unified centralist (unitarian) government or whether it should be ruled by decentralised governments in the socialist republics within the Yugoslav Federation. Legally, a large decentralisation of power from federative to republican level was introduced in the 1974 Constitution,

---

<sup>48</sup> Pranjković uses the term 'socialist vocabulary' instead of 'federal', introduced earlier (p.11) as Brozović's term

<sup>49</sup> In The Croatian National Corpus the ratio for the federal words 'muzika' and 'sistem'/'sistema' towards their Croatian counterparts 'glazba' and 'sustav' are: 1:45 and 1:35, resp.

but at the same time there was a crack-down on all political opposition to the party line, often labelled as nationalism or liberalism. (Mønnesland 1999: 222–28). So, the Party leadership in the republics gained more power over their “own” republic, but at the same time the central control over these Party leaders was strengthened. Therefore, I came as no surprise to citizens in the former Yugoslavia that Croatian was legally and officially established as a separate language shortly after the state of Croatia declared its independence in 1991. However, it took some time before international scholars of Serbo-Croatian began using the new lingonym – Croatian.

### **1.1.6 Partition - Independency**

If we accept the emic classification (based on the symbolic and social status of language in the minds of their speakers) and the authority of the official naming of a language (in legislation etc.), 1991 is the point in time where the relatively short history of standard Serbo-Croatian comes to an end and where Serbian and Croatian part and become officially independent of each other.

If we were to look upon the linguistic system (the grammar) of Serbian and Croatian there is no doubt that the two share a common linguistic system. If we regard the linguistic system as the defining factor Serbian, Croatian, Bosnian and Montenegrin are surely one language. There are, however, differences in the orthographic representation which, in turn, have nothing to do with the underlying linguistic system. There are differences in the lexis, the affixes and in the morphology and syntax which do not represent systemic differences either. What *is* different are the norms and the names of the languages. Here, I distinguish between two kinds of norm: The norm that is prescribed by language authorities and the norm which govern the usage of the language, which I shall call *the prescribed norm* and *the usage norm*.

As opposed to naming the unified language, there was never a discussion of how to designate the language in Croatia and Serbia after the break-up of Yugoslavia. As we saw above, the declaration of a separate Croatian language, which followed the declaration of independent Croatian state, did not appear out of the blue, but had been underway for some time and the fact that distinctly Croatian expressions, vocabulary and constructions had been labelled ‘nationalistic’ by the former regime only lent them even more prestige in a situation in which nationalism was word with distinctly positive, patriotic connotations as opposed to its distinctly negative, traitorous connotations during the former communist regime. Even though ‘nationalistic’ tendencies had

been emerging in the Croatian variant of Serbo-Croatian before 1991, after 1991 radical changes occurred, especially in the vocabulary norm. The prescribed norm changed very rapidly and speakers of Croatian were eager to implement these changes, so some changes quickly became part of the usage norm as well. As the renowned Croat normativist Stjepan Babić puts it: 'In 1991, once again the process of 'croatisizing' Croatian began' (Babić 2009: 176)<sup>50</sup>.

In Serbia the change in name from Serbo-Croatian to Serbian happened in 1992 when, in article 15 of the new Constitution for the Federal Republic of Yugoslavia, consisting of Serbia and Montenegro, it is stated that the official language is Serbian in both its Ekavian and Ijekavian version.<sup>51</sup> However, the change of name and what led to it did not cause the Serbian standard language to change, besides in name, when one compares the language in the 80's with the 90's (Radovanović 2009: 199). No radical changes happened in grammar, lexis or orthography (Bugar-ski 2009: 134). There were no attempts from the authorities themselves or backed by them at changing the language, i.e. altering the prescribed norm and thereby the usage norm into something "more Serbian" now that it had become the language of Serbs and not of Croats, Bosniacs etc. (Ibid: 135). Serbian lexicographers still in general regarded Serbian and Croatian as one language and therefore continued to include lexemes which during the Serbo-Croatian period were regarded as part of the western (i.e. Croatian) variant of Serbo-Croatian. (Stijović 2009: 219–20)

The linguistic system of Croatian did not change either. What did change were the prescribed norm and the usage norm. Both the prescribed norm and the usage norm of the Croatian language changed through the promotion of what was deemed more Croatian and the demotion if not banishment of what was considered not to be 'in the spirit of the Croatian language'. The aim was to (re)discover what was purely Croatian. This was partly done by trying to eradicate Serbian and other non-Croatian influences in the language, primarily in the lexis (Langston and Peti-Stantić 2003: 247, 255). So, in contrast to Serbian, radical changes in the usage of existing lexis, combined with the invention of new lexis did happen in Croatian. The changes in the accepted usage norm were so sudden that public figures developed what

---

<sup>50</sup> Original: „...1991 beginnt man die kroatische Sprache wieder zu kroatisieren.“

<sup>51</sup> Original: Član 15. U Saveznoj Republici Jugoslaviji u službenoj upotrebi je srpski jezik ekavskog i ijekavskog izgovora i ćirilčno pismo, a latiničko pismo je u službenoj upotrebi u skladu sa ustavom i zakonom.

has been termed a 'fear of the mother standard tongue'<sup>52</sup> and thus lost their eloquence when speaking in public for fear of using words that had so quickly become unacceptable. Whether this situation is best described as the Croats finally being able to speak and use a pure Croatian again after more than a century of oppression or, as a vigorous attempt to diminish all that reminded the Croatian speakers of Serbia and the union with Serbia or, as a reflection of the fact that Serbo-Croatian had never existed does not change the fact that the prescribed norm and the usage norm (was) changed.

Language is seen as a symbol of national identity and with the establishment of an independent Croatian state, independent of Serbia, a lot of effort was put into making the Croatian language independent of Serbian. The discussions between linguists both within Croatia and between Croatia and Serbia have been, to say the least, heated. In the polemics between otherwise respected linguists and other scholars personal attacks have been made where even the opponent's mental health has been questioned not to mention accusations of poor academic knowledge. Cf. Pranjković (1997; 2008) and specifically for conflicts within Croatia: Babić and Ham (2005: 22-30-65-109-148). Today in Croatia, conflicting views on orthographic issues still linger even though the Ministry of Science, Education and Sport in July 2013 recommended one particular orthographic manual<sup>53</sup>, in accordance with which textbooks for primary and secondary schools Croatia are now written.

In Serbia, the Ministry of Education (in 1997) and The Board for Standardisation of the Serbian Language (in 1999) recommended an Ekavian orthographic dictionary, published by the Matica Srpska, whose first edition was published in 1993, altered and supplemented in 2010, the most recent one being from 2013<sup>54</sup> with an Ijekavian edition published in 2014 This orthographic dictionary enjoys the position of being the official orthographic manual of Serbian in Serbia and in Republika Srpska<sup>55</sup> (Srpski jezički atelje 2014). As in Croatia, other, competing, orthographic dictionaries of various quality have been published in Serbia, none of which have been recommended by the authorities (V. Brborić 2008: 52).

---

<sup>52</sup> Original: "strah od materinskoga standardnog jezika" coined by Nives Opačić

<sup>53</sup> Hrvatski pravopis Instituta za hrvatski jezik i jezikoslovlje

<sup>54</sup> Pešikan, Jerković and Pižurica (2013) *Pravopis srpskoga jezika*, Matica Srpska – Novi Sad

<sup>55</sup> The Serb lead part of Bosnia-Herzegovina

In later years, after the lingonyms Croatian and Serbian have stabilized in usage, focus among linguists and normativists from both speech communities have turned upon the so-called ‘onslaught of English words’, not at least in the language of the media. The focus, in other words, is mostly on the lexical changes that have come about in Croatian *and* Serbian due to contact with English, but changes in word-formational patterns, construed as being induced by contact with English, have also been noted, and in both speech communities efforts have been done to curtail this development of the usage norm. In Croatia, in order to counteract the so-called an Anglification of Croatian, competitions are announced to find the best substitution of foreign words and among the criteria in such a competition announced by the prestigious linguistic periodical *Jezik* it is emphasised that the focus should be on substituting English loanwords (Anglicisms) (Ham 2010: 115)<sup>56</sup> The Institute for Croatian Language and Linguistics<sup>57</sup> has also reacted on the influx of Anglicisms and seeks to animate the public in a similar way by asking for suggestions for substitutions of loanwords via their project *Bolje je hrvatski* (‘It is better (in) Croatian’) at [www.bolje.hr](http://www.bolje.hr). The project began in March 2015 and highlights 5 suggested substitutions per month hitherto (July 2015) all highlighted suggestions are substitutions for English terms (‘Bolje je hrvatski’ 2015). The influx of so-called unnecessary loans from English into both Croatian and Serbian are besides being seen as an effect caused by globalisation also seen as a sign of the lack of concern for properness and a tendency to be ‘seduced by fashion’ and in ‘widespread spiritual laziness and conformism’ (Opačić 2007: 24; T. Prčić 2006: 419). And, according to Prčić (2006), the English words and constructions are put into use according to the principle of ‘Write it as you want, pronounce it as you like, the grammar is up to you and the meaning of it is haphazard’.(T. Prčić 2006: 417)<sup>58</sup>

So, despite the different attitudes towards the former common language of Serbo-Croatian and the high awareness in Croatia of what is acceptable Croatian and what is not, Serbian and Croatian seem to be facing similar challenges in terms of new linguistic behaviour induced by

---

<sup>56</sup> The lexical divergence between Serbian and Croatian concerning replicated will be described, exemplified and discussed further in section 4.1.

<sup>57</sup> The Institute, which was responsible for the compromise-seeking orthographic dictionary, recommended by the Croatian government (mentioned above) is a Government-financed body, which is considered to be more moderate in its views on language than the editorial line in the periodical *Jezik*.

<sup>58</sup> Nives Opačić describes Croatian and Tvrtko Prčić describes Serbian.



contact with English. Another common trait is the, to some extent, similar way in which the linguistic authorities of the two languages respond when faced with the linguistic effects of globalisation.

## 1.2 Differences

Whenever we describe a language, for instance my own mother-tongue Danish, we cannot subsume all that is defined as Danish. What we do describe is a particular variety of Danish. The label “Danish language” subsumes all varieties, be they geographically defined dialects or socially defined sociolects or even stylistically defined genres or functional styles. In other words; a survey of how youngsters speak Danish in a particular part of society is a survey of Danish but an analysis of the usage of a particular adverb in 19<sup>th</sup> century Danish poetry is also a survey of Danish. Among the many varieties of Danish as of many other modern languages, there is one variety which has been promoted to a nationwide level; this variety of Danish is, in modern terms, the Danish standard language or standard variety of Danish.<sup>59</sup>

The difference between variational difference in Danish and the variational difference between two specific varieties of Serbo-Croatian lies in the fact that not one but two varieties were promoted to the status of standard variety which, if recognised as such by scholars, were called variants of the same standard language. These variants were, as we have seen, later promoted to first two, then four standard languages.

Most standard languages are planned. Serbo-Croatian, Serbian and Croatian are no exception. In the following chapter we will return to what language planning means and what it meant for Serbo-Croatian. At this point it suffices to say, as did the renowned Croatian linguist and lexicographer Vladimir Anić, that Serbo-Croatian could not be planned according to the utopian ideal of: *one people – one language – one nation/ethnos – one script – one orthography – one pronunciation – one vocabulary* (Anić 1990: 101). No other language can fully live up to this ideal but Serbo-Croatian does not fulfil *any* of the notions in the list, except *maybe* one language. If the other notions, i.e. one nation/ethnos – one script – one pronunciation – one vocabulary are preconditions for the existence of one language, Serbo-Croatian is definitely *not* one language. In other words, in order for standard Serbo-Croatian to live up to this ideal, Ijekavian or Ekavian should have been excluded, Cyrillic or Latin script should have been excluded, all variation in orthographic rules should

---

<sup>59</sup> In Danish: *rigsdansk* ('Danish of the Realm')

have been eliminated and finally all true synonyms should be used by all speakers of standard Serbo-Croatian or alternatively, one lexeme should be granted the status of superregional as opposed to its synonyms, which in turn would be stylistically or regionally marked, and thus cease to be true synonyms of the superregional lexeme. Finally, Serbs and Croats should have embraced the idea of belonging to the same nation or ethnos.

In the planning of Serbo-Croatian as one language a different and much more flexible approach was taken. This led to normative works on Serbo-Croatian, which listed differing rules and regularities for the same word-formational, grammatical, phonological and orthographic phenomena depending on script, reflex of *jat*, spelling of analytical future tense, declension of certain nouns and possessive pronouns, even punctuation. Very often this difference in rules and regularities could be ascribed to dialectal, regional and/or national differences. In Vukovian normative works on Serbo-Croatian in both the interwar period and in the period following the Novi Sad Agreement (1954) the national and/or regional origin of the different rules and regularities were, deliberately, rarely mentioned. Later, following the general acceptance among scholars of the existence of at least two, at most four, variants of Serbo-Croatian, normative work was done under the heading of Croatian, Serbian or Serbo-Croatian.

As mentioned above, the term variant was introduced into the scholarly circles in 1965 by Milka Ivić when she addressed the issue of norm in standard Serbo-Croatian. The main precondition for the existence of a variant (as opposed to a variety) is that civilizational development of the base standard has taken place. And this development *had* happened in the case of the Serbian, Croatian, Bosnia-Herzegovinian and Montenegrin manifestations of the base standard, as Dalibor Brozović (leading Croatian linguist) puts in his 1985 article *On the functioning of languages in the Yugoslav Federation*<sup>60</sup> (1985: 84). Even Pavle Ivić (leading Serbian linguist) in several publications (1968, 1971, 1986 and 1990) calls attention to the fact that variants do exist, have existed since the integration of Serbian and Croatian in the mid-19<sup>th</sup> century, and are the product of different civilizational developments within the Serbo-Croatian area, especially around the two cities of Zagreb and Belgrade (1990 [1968]: 314), (1986 [1971]: 193), (1990 [1986]: 127) (1990 [1989]: 323). So, an entirely unified Serbo-Croatian has never existed. Interest-

---

<sup>60</sup> „O funkcioniranju jezika u jugoslavenskoj federaciji“

ingly, variance defined by the existence of different non-linguistic differences such as “civilizational development” is and have been recognised throughout the history of Serbo-Croatian.

The linguistic and conventional differences between modern standard Serbian and Croatian (variants/languages) concern lexical elements (words, semiwords and affixes), gender assignment on specific nouns, syntactic constructions, spelling rules, accentuation, pronunciation and punctuation.

### 1.2.1 Scripts, pronunciation and spelling rules

Standard Serbo-Croatian and standard Serbian may be written in both the Cyrillic and the Latin script. Standard Croatian may only be written in the Latin script.

*Different rules for transliteration, depending on script and language:*

<u>Cyrillic (Serbian)</u>	<u>Latin (Serbian)</u>	<u>Latin (Croatian)</u>
Дума (<Dumas) or Dumas(<Dumas)	Dima(<Dumas)	Dumas (<Dumas)
Пјеп (<Pierre)	Pjer (<Pierre)	Pierre (<Pierre) or Pierre (<Pierre)

Serbo-Croatian orthographic manual and dictionary (Hraste et al. 1960: 144-154). Serbian orthographic manual and dictionary (Pešikan et al. 1993: 186). Croatian orthographic manual and dictionary (Babić et al. 2008: 41)

*Difference in spelling rules when the future tense is written VERB + AUX*

#### Serbian

*pjevaćemo* ('we shall sing'): synthetic future consisting of a stem of the infinitive *pjevati* (*to sing*), i.e. *pjeva-* and the enclitic present of the auxiliary *htjeti*, i.e. *-ćemo*

#### Croatian:

*pjevat ćemo* ('we shall sing'): analytic future consisting of a clipped version of the infinitive *pjevati* (*to sing*), i.e. *pjevat* and the enclitic present of the auxiliary *htjeti*, i.e. *ćemo* (Stevanović 1964: 1: 362).

Stevanović, characteristically, does not mention that the former spelling is used in the Serbian or Eastern variant, whereas the latter is the norm in the Croatian or Western variant. Today it is one of the often mentioned differences between Serbian and Croatian, e.g. in (Mønnesland 2002: 39)

and in (B. Brborić 2011: 20). A recent corpus-based comparison of Croatian, Serbian and Bosnian confirms this difference (Bekavac et al. 2008: 34–35). There is, however, no difference in the pronunciation (/’pjɛvaʦɛmɔ/) or in the word order rules governing the future tense. Therefore, in my view, the differentiation in categorisation is not necessary. Instead, it should be viewed as one type of future tense, spelled in two different ways. It is, in other words, merely a conventional difference, not a linguistic one.

*Differences in pronunciation of h /x/ and syllable-final l /l/*

Stevanović (1964: 1: 138–40) devotes almost three pages to *h* /x/ in which he explains how it is due to the usage in the south-western parts of the Štokavian dialectal continuum (Dubrovnik and southwards along the coastline of today’s Montenegro) as well as in the usage among Muslims in Bosnia and the usage by Non-štokavian Croats, that *h* exists in standard Serbo-Croatian orthography.

Today the different pronunciation and spelling due to the presence or absence of /x/ is used as a differential marker between standard Croatian and standard Serbian:

<u>Croatian</u>	<u>Serbian</u>	
<i>kuhati</i>	<i>kuvati</i>	‘to cook’
<i>suh</i>	<i>suv</i>	‘dry’
<i>snaha</i>	<i>snaja/snaha</i>	‘daughter-in-law’
<i>hrđa</i>	<i>rđa</i>	‘rust’

(Langston and Peti-Stantić 2014: 61)

In Štokavian, syllable-final /l/ in word-final position or preceding a consonant vocalizes into /o/, which is why wordforms whose stem ends in /-l/ will end in /-o/ if no ending is added to the stem. For example, the perfect participle, feminine, singular consists of the verbal stem of the infinitive form + the participle suffix: /-l/ + an inflectional suffix marking number and gender: /-a/. This we see, for instance, in the perfect participle active *bila*, formed of the infinitive stem of *biti* (‘to be’): *bi-*, the participle affix *-l* + and the inflectional affix *-a*.

Serbian and Croatian

bi-	-l-	-a	> <i>bila</i>
INF. STEM	PTC. AFFIX	AFFIX marking GENDER, NUMBER	PERF.PTC.FEM.SG

*Ona je bila* – ‘She was’

bi-	-l-	-Ø	> <i>bio</i>
INF. STEM	PTC. AFFIX	AFFIX marking GENDER, NUMBER	PERF.PTC.MASC.SG

*On je bio* – ‘He was’

When the ending that marks gender and number is a zero-morph (-Ø), as in the masculine, singular, the affix /-l/ becomes syllable-final and thus vocalizes into /o/: \**bil* > *bio*. This rule applies systematically (almost) independently of word category in both the Serbian and Croatian standard language. There is, however, in standard Croatian a tendency to retain a syllable final /l/ in nouns in which the preceding vowel is /o/. Thus we have the following differences between standard Croatian and standard Serbian.

Croatian	Serbian	
<i>stol</i> (< <i>stol</i> )	<i>sto</i> (< <i>stol</i> )	‘table’
<i>vol</i> (< <i>vol</i> )	<i>vo</i> (< <i>vol</i> )	‘ox’

(Langston and Peti-Stantić 2014: 61)

The syllable-final /l/ is as is the pronunciation of the phoneme /x/ a differential marker between standard Croatian and standard Serbian.

### ***The fate of the Cyrillic script***

In a statement concerning a proposed amendment to ‘The Law on Official Use of Language and Scripts in the Serbian Autonomous Province of Vojvodina’ made by ‘The Board for Standardisation of the Serbian Language’, founded 1998, this board emphasises the fact that the Latin alphabet was introduced and won leeway among Serbs following the establishment of the first Yugoslavia and that this script slowly but surely suppressed the Cyrillic one, decidedly so in Communist Yugoslavia, where ‘the authorities systematically favoured both public and administrative use of the Latin alphabet’<sup>61</sup> (Brborić et al 2006: 129). In time, this led to the near abandonment of the Cyrillic alphabet in Croatia, Bosnia-Herzegovina and even in the Vojvodina province of Serbia (Ibid).

The tendency to forego the Cyrillic alphabet is still strong, even though it was proclaimed the official script in Serbia and Montenegro in the 1992 Constitution of The Federal Republic of Yugoslavia, which The Standardisation Board, in a later statement, ascribes to the presence of

---

<sup>61</sup> Original: “власт је систематски фаворизовала јавну и службену употребу латинице”

old and new globalist prejudices<sup>62</sup> meaning that not only the unification policy during communist time sought to suppress Cyrillic but the globalisation, fronted by English and therefore the Latin alphabet, leads the people to reinforce and continue this suppression (Brborić et al 2006: 170).

Today, the Cyrillic script has a very strong symbolic value, even more so among Serbs living in Croatia as events in September 2013 in Vukovar, which has a large minority of Serbs, have shown. Bilingual (Serbian-Cyrillic and Croatian-Latin) signs were put on governmental and municipal buildings in accordance with an EU-motivated minority rights law<sup>63</sup> that makes bilingual signs mandatory in any area where more than one-third of population belongs to an ethnic minority. This, in turn, led to aggressive reactions among Croats, especially the powerful War Veterans Organisation. There were incidents of vandalism, riots and even demands of a referendum to get rid of the Serbian, Cyrillic script from signs in Croatia.<sup>64</sup>

In Serbia, today, the Cyrillic script is still in use and is taught in schools alongside the Latin script. A “Law on Official use of Languages and Scripts” introduced in 1991 and amended last in 2010 prescribes the primary use of Cyrillic but allows the use of Latin script, either as a supplement to the mandatory Cyrillic text and in some instances as a mandatory doublet (signs on international and main roads). ‘Official use’, however, far from covers all written communication<sup>65</sup>. It is not mandatory that text books, school books, other books, newspapers, magazines, shop signs, not to speak of electronic media are written in the Cyrillic script. It would be interesting to learn how many and which publications in Serbia are in Cyrillic. Unfortunately, The Statistical Office of the Repub-

---

<sup>62</sup> Original: „У бившој СФРЈ, од самог њеног настанка (1945. г. – под називом ФНРЈ, важећем до 1963), латиницу је подупирала владајућа комунистичка идеологија, а сада је, чини се, подупире одсутност једнозначне идеологије и присутност старих и нових предрасуда, мондијалистичких, глобалистичких и других.”

<sup>63</sup> Croatia joined the EU on July 1, 2013

<sup>64</sup> <http://www.bbc.com/news/world-europe-23934098> and <http://www.balkaninsight.com/en/article/vukovar-bilingualism-introduce-faces-violent-resistance> (Accessed 2015, 7/10)

<sup>65</sup> “Official Use” only pertains to communication from public and among authorities, keeping statutory records, public documents, documents of labour-related rights and duties and responsibilities of employees and finally in public announcements and indications (signs)(Official Gazette of RS, No. 30/2010, Art. 3)

lic of Serbia, which has statistical information on all publications in Serbia, does not register which alphabets the publications are written in.<sup>66</sup> The lack of statistical information might be an expression of the unawareness or unwillingness to counteract the claimed suppression of Cyrillic, mentioned at the beginning of this section but such a claim would be, I admit, a purely speculative one.

### 1.2.2 Lexemes, semiwords and affixes

The standard Croatian norm prescribes usage of domestic, i.e. Croatian, vocabulary whenever possible, whereas standard Serbian norm does not. This difference in attitude towards lexical items of foreign origin results in different prescribed vocabulary.

#### *Different prescribed vocabulary*

<u>Serbian</u>	<u>Croatian</u>	
univerzitet	sveučilište	'university'
fudbal	nogomet	'football'
muzika	glazba	'music'
sekretarica	tajnica	'secretary'
internet	međumrežje	'internet'
faktor	čimbenik	'factor'
ekonomija	gospodarstvo	'economy'
sala	dvorana	'hall'
direktor	ravnatelj	'director'
režiser	redatelj	'film director'
ambasador	veleposlanik	'ambassador'
oficir	časnik	'military officer'
štampa	tisak	'(the) press'
lingvista	jezikoslovac	'linguist'
geografija	zemljopis	'geography'
Ujedinjene Nacije	Ujedinjeni Narodi	'United Nations'
<u>sajber</u> -kriminal	<u>kiber</u> kriminal	'cyber-crime'
<u>multinacionalan</u>	<u>višenacionalan</u>	'multinational'

---

<sup>66</sup> This information was obtained July 13, 2015 through correspondence with Ms Sunčica Stefanović Šestić, Head of statistics on education, science, culture and other societal businesses in The Dissemination and Public Relations Division.

*Different adaptation of foreign lexical matter (lexemes)*

<u>Serbian</u>	<u>Croatian</u>	
finansijski	financijski	'financial'
demokratija	demokracija	'democracy'
minut	minuta	'minute'
leksema	leksem	'lexeme'
kriterijum	kriterij	'criterion'
lingvista	lingvist	'linguist'
hemija	kemija	'chemistry'
manipulisati	manipulirati	'to manipulate'
organizovati	organizirati	'to organise'
restartovati	restartati	'to restart'

*Different word-formational (derivational) regularities*

Serbian, Serbo-Croatian

*čitalac* ('a reader'): a nomen agentis formed from the verbal stem čita- (read) and the suffix -lac

Croatian

*čitatelj* ('a reader'): a nomen agentis formed from the verbal stem čita- (read) and the suffix -telj.

By the Vukovites the suffix *-telj* is described as archaic and/or due to Russian influence (Stevanović 1964<sup>1</sup>: 540), and Maretić (1963 [1899]) deals with this suffix in a similar way adding that there are very few *nomens agentis* derived in this manner (Maretić 1963 [1899]: 362). Today, the suffix *-telj* is frequently mentioned as typical for Croatian and atypical for Serbian (Brborić 2011: 21; Langston and Peti-Stantić 2014: 63–64; Bekavac et al. 2008: 35). Other affixes, marked either as Croatian or Serbian are e.g. *su-/sa-*, *ero/oro*, *ić/će*:

<u>Serbian</u>	<u>Croatian</u>	
<i>savremen</i>	<i>suvremen</i>	'contemporary'
<i>petoro</i>	<i>petero</i>	'five'
<i>ormarče</i>	<i>ormarić</i>	'small cabinet'

(Langston and Peti-Stantić 2014: 63–64).

This does not mean that the affixes do not exist in both standard languages, only that one is preferred over the other.



### 1.2.3 Morphosyntactic differences

The most commonly emphasised syntactic difference between Serbian and Croatian concerns the possibility in Serbian as well as Croatian to place either an infinitive or a dependent clause with a verb in the present tense introduced by the conjunction *da* as the complement of different types of heads, the most common ones being, modal verbs, phasal verbs, verbs of motion, nouns and not at least main verb in the unmarked future tense, Future I.

#### Modal

*Moram da idem*  
I must CONJ I go-PRES

*Moram ići*  
I must go-INF

‘I have to go.’

#### Phasal:

*Počinem da čitam*  
I begin CONJ I read-PRES

*Počinem čitati*  
I begin read-INF

‘I’m beginning to read.’

#### Motion:

*Idem da spavam*  
I go CONJ I sleep-PRES

*Idem spavati*  
I go sleep-INF

‘I’m going to bed.’  
(i.e. so that I may

#### Noun:

*On ima običaj da pjeva*  
He has habit CONJ sings-PRES

*On ima običaj pjevati*  
He has habit sing-INF

‘He usually sings.’

#### Future I:

*Ja ću da dođem*  
I will CONJ I come-PRES

*Ja ću doći*  
I will come-INF

‘I’ll come.’

(Mønnesland 2002: 204–5)

In standard Croatian, the infinitive is in these cases strongly preferred (Langston and Peti-Stantić 2014: 65) whereas both constructions are acceptable in Serbian standard depending on which functional style of the standard language is in question (Piper and Klajn 2013: 412). For instance, the *da*-construction in Future I is not recommended in writing by Serbian normativists (Piper and Klajn 2013: 173).

The dependent infinitive may also substitute the *da*-construction even though the subject of the dependent clause is an object of the verb in the main clause:

*Ona me je naučila da radim*  
She me has taught CONJ I work-PRES

*Ona me je naučila raditi*  
She me has taught work-INF

‘She taught me (how) to work.’

*On mi je dozvolio da uđem.*  
He me has allowed CONJ I enter-PRES

*On mi je dozvolio ući.*  
He me has allowed enter-INF

‘He allowed me to enter.’

In these instances the *da*-construction is in frequent use, also in the Croatian norm (Mønnesland 2002: 204–5). The difference between the two

constructions, i.e. *infinitive* vs. *da*-construction, is undoubtedly a syntactic one, but it is not an absolute difference between Serbian and Croatian despite what many (Croatian) linguists claim because the difference lies not in the linguistic system. The difference lies primarily in the prescribed norm and partly in the usage norm.

#### *Morphological differences*

In declension of certain masculine proper nouns ending in –o, the Serbian and Croatian standard differ in respect to which of the nominal declensions the nouns should follow:

	<u>Serbian</u>	<u>Croatian</u>
Nom:	Ivo	Ivo
Voc:	Ivo	Ivo
Acc:	Iva	Ivu
Gen:	Iva	Ive
Dat:	Ivu	Ivi
Instr:	Ivom	Ivom
Loc:	Ivu	Ivi

Possessive pronouns in 3<sup>rd</sup> person (*njegov* ‘his’, *njen/njezin* ‘her’ and *njihov* ‘their’) may according to the Croatian standard only be declined according to the indefinite declension, whereas it is frequently declined according to definite declension in Serbian:

	<u>Indefinite</u> <u>“Serbian”</u>	<u>Definite</u> <u>“Croatian”</u>
Nom/Acc/Voc, masc., sing.:	njegov	njegov
Gen, masc., sing.:	njegovog	njegova
Dat/Loc, masc., sing.:	njegovom	njegovu
Instr, masc., sing.:	njegovim	njegovim

For the sake of clarity I will, once again, emphasise that the differences listed in subsections 1.2.2-1.2.3 are not absolute differences, but rather represent tendencies and preferences which are not only expressed through the usage norm but also through the prescribed norm, formalized in various normative works and therefore promoted in society. As I am interested in contact-induced linguistic changes which may enhance or diminish the differences between Serbian and Croatian, I will investigate further some, but not all, of the differences mentioned above.

### 1.3 The linguistic phenomena under investigation

The Serbian sociolinguist and language planning theoretician Milorad Radovanović has on several occasions pointed out what he sees as the main language innovations in modern day Serbian. Among the language changes in Serbian which are supposedly triggered by language contact we find the high rate of replicated lexemes, especially from English (of which the nouns are far the most numerous). Secondly, an increase in nominal compounds consisting of an undeclined noun immediately followed by another noun, viz. (N<sub>UNDECL</sub>N)-constructions has been observed. Thirdly, a high productivity in word-formation with semiwords<sup>67</sup> as well as with replicated affixes and fourthly an increased use of the numeral *jedan* and the pronoun *neki* as indefinite articles (Serbian and Croatian are defined as article-less languages) (Radovanović 2000: 28; 2009: 210–11). The same phenomena have been identified in Croatian by different Croatian linguists (Skelin Horvat 2004; Starčević 2006; Pranjković 2000; Barić 1980; Mihaljević 2012)

In the investigation and analysis of empirical data as well as of the existing research on these innovations in both Croatian and Serbian, I will focus on these four phenomena and aim to ascertain whether or not Serbian and Croatian are developing in the same manner or in different manners.

#### 1.3.1 Nouns

Examples: *tržništvo/marketing* (<'marketing'), *pokazivač/kursor* (<'cursor'), *naraštaj/generacija* (<'generation'), *kakvoća/kvalitet(a)* (<'quality'), *ugodaj/štimung* (<'Stimmung')

#### 1.3.2 Nominal compounds

Examples: *pop pevač* ('pop singer'), *seks bomba* ('sex bomb'), *šok terapija* ('shock therapy'), *biznis klasa* ('business class'), *menadžment sposobnosti* ('management skills'), *nana čaj* ('mint tea'), *internet veza* ('Internet connection')

#### 1.3.3 Semiwords and affixes

Examples: *papirologija* ('excessive amounts of paper'), *bubnologija* ('expertise in drum-playing'), *kobasicijada* ('sausage competition'), *ćutolog* ('quiet, introvert person'), *tračér* ('gossiper, i.e. person prone to gossip'), *kavoman* ('coffee-maniac', i.e. 'person obsessed with coffee'),

---

<sup>67</sup> In Serbian and Croatian terminology: *prefiksoidi* ('prefixoids') and *sufiksoidi* ('suffixoids')

*reizbor* ('reappointment'), *aeroprostor* ('air space'), *akvapromet* ('water traffic'), *hipersila* ('hyper-power')

### 1.3.4 Indefinite articles

The numeral *jedan* and the indefinite pronoun *neki* functioning as indefinite articles:

*Povremeno bi mu i neki prijatelj donio litru domaće rakije.*

'From time to time, a friend brought him a litre of homemade brandy.'

*Nekoliko dana prije toga od jednog prijatelja u Münchenu čuo sam kako...*

'A couple of days earlier, I heard from a friend in Munich that..'

## 1.4 The empirical data

In order to gain insight into the usage norm of Croatian and Serbian I have turned to two public electronic text corpora. A text corpus of contemporary Croatian *Hrvatski Nacionalni Korpus* ('The Croatian National Corpus'), abbreviated HNK which holds approx. 100.8 million words and the *Korpus savremenog srpskog jezika* ('Corpus of Contemporary Serbian', abbreviated SrpKor) which holds approx. 122.2 million words. Both corpora consist primarily of text from the media but also contain text from other genres.

The language of the media is a particularly good representative of language usage which is in accordance with the *usage norm*. The representativeness of the media as empirical data for research into linguistic innovations is underlined by authorities on language policy (Joseph 2006: 117), (Busch 2010: 190) (Sandøy and Selback 2007: 11) and scholars of Serbian (Stanković 1997: 87), Croatian (Hudeček and Mihaljević 2009: 7), (Mihaljević 2002: 228), (Skelin Horvat 2004: 98), (Silić 2006: 92–93) and Serbo-Croatian (P. Ivić 1992: 102).

### *Functional styles*

According to both Serbian and Croatian tradition, language usage is usually divided into five different functional styles: The Belles-lettres or Literary style (in the genre of creative writing), The Publicistic or Publishing style (the language of the written media), The Scientific style (in the genre of scholarly dissemination), The Administrative or Bureaucratic style<sup>68</sup> (of official documents and papers) and The Conversational or Colloquial style (covering both oral and written conversation) (Tošović 2002; Silić 2006).

---

<sup>68</sup> In English terminology also called The Official Style

Stanković (1997) refers to a long tradition of research into the differing functional styles of language and claims along the lines of Mihaljević, Ivić, Joseph and Busch that the publicistic style (the written language of the media) is where you will find the implementation of the newest developing tendencies in a language, both because media report on and reflect the current events in society (social, financial, cultural etc.) but also because they are quick to implement the linguistic innovations that occur among speakers.

Furthermore, the media are quick to pick up on linguistic innovations occurring in other functional styles (administrative, literary, and scientific), thus circulating these innovations. Through the media, innovations which originate in a specific area or among a specific group of people very quickly reach a large audience and thus enhances the chances for an innovation to be implemented by a large number of speakers, which may normalise the innovation, i.e. make it part of the usage norm. So, the media is not only the best place to search for linguistic innovations, it is also the best place to promote an innovation.

#### *Corpus of Contemporary Serbian – Croatian National Corpus*

The first version of the *Corpus of Contemporary Serbian* (SrpKor) was completed in 2003 and was funded by the Serbian Ministry of Science and placed in the Hands of the Human Language Technology Group at the Department of Informatics and Computer Science, Faculty of Mathematics, University of Belgrade where it still resides. The project is led by Duško Vitas and developed and maintained by Miloš Utvić.

The *Croatian National Corpus* (HNK) was initiated in 1998 and financed by The Ministry of Science and Technology in Croatia and subsequently supported through nationally funded projects. (Tadić 2009: 220). Since 2011 it is also funded by EU. HNK resides at Institute of Linguistics, Faculty of Humanities and Social Sciences, University of Zagreb. The primary researcher is Marko Tadić, Head of the Chair for Algebraic and Computational Linguistics at University of Zagreb. (CESAR 2013: 182)

#### *The sources in HNK*

The HNK-project has, since 2004, exclusively been focusing on Croatian text from 1990 and later. (Ibid: 221).<sup>69</sup> The version of the HNK used in this dissertation (HNK 2.5) is from 2009 and contains approx. 100.8 m.

---

<sup>69</sup> Curiously, Tadić (2009: 221) counterfactually claims that earlier dated samples were removed from the HNK v. 2.5.

wordforms or *word tokens*<sup>70</sup>. Whereas a *word token* stands for any word-form in the corpus, a *word type* stands for all instances of the same word token, i.e. the number of *different* word tokens.<sup>71</sup> However, homographic word types are counted as one word type, which makes the probable number of different word tokens higher than that of word types.

A newer version from 2013, the HNK 3.0, also exists and it contains 216.8m word tokens. It is, however, scarcely described and documented, so I have not found it possible to establish its representativeness. This is however not the case with the 2.5 version which has been online since April 2009 and is compiled primarily of texts published between 1990 and 2005 but also contains 65 Croatian classic literary works which date primarily from the late 19<sup>th</sup> century and early 20<sup>th</sup> century and a few from the 16<sup>th</sup>, 17<sup>th</sup> and 18<sup>th</sup> centuries.<sup>72</sup> These Croatian classics comprise of 3.6m word tokens (i.e. approx. 3.7% of the HNK 2.5). The HNK aspires to the ideal structure of 74% informative (factual) texts, 23% fictional texts and 3% mixed texts<sup>73</sup>. The factual texts are collected from newspapers, magazines, scientific journals, books, brochures and similar, whereas the fictional texts consist of novels, short stories, diaries and similar. The documents and texts in HNK 2.5 have been collected from the following newspapers, magazines, journals and fictional literature:

*Daily newspapers:*

„Dubrovački vjesnik“, 2005, approx. 0.8m word tokens.

„Večernji list“, 1999, approx. 2.2m word tokens.

„Vjesnik“, 2000-2003, approx. 46.6m word tokens.

„Glas Slavonije“, 2002-2005, approx. 17m word tokens.

---

<sup>70</sup> The number of word types was reached by extracting a word list of all wordforms in HNK 2.5 which resulted in 11418 pages of 101 word types and one page of 8 word types. I have extracted the number myself because there is some confusion in the literature as to the amount of tokens and word tokens, respectively in HNK 2.5. According to Tadić (2009: 221) the number of tokens was then approximately 104.3m, while (CESAR 2013: 182), referring to information updated in 2011, informs us that the number is 101m tokens. According to Dobrić (2012) the number of word tokens in the HNK 2.5 is 101.3m. The HNK v.2.5 which I have accessed during my research holds exactly: 100,884,284 word tokens ([http://filip.ffzg.hr/bonito2/run.cgi/sub-corp?corpname=HNK\\_v25](http://filip.ffzg.hr/bonito2/run.cgi/sub-corp?corpname=HNK_v25), Accessed January 28, 2016) so it appears that the information in CESAR (2013) is the most accurate of the cited.

<sup>71</sup> A *word type* should not be confused with a lemma, as a lemma is the abstract representative of a lexeme, which may manifest in several word types, especially in languages with a rich flexive morphology such as Croatian and Serbian.

<sup>72</sup> For a complete list of these works, see: [http://www.hnk.ffzg.hr/Izvori\\_Klasici.html](http://www.hnk.ffzg.hr/Izvori_Klasici.html) (Accessed Oct 10, 2016)

<sup>73</sup> <http://hnk.ffzg.hr/struktura.html> (Accessed July 13, 2015)

#### *Weekly news magazines*

“Croatia Weekly” 1998-2000, approx. 1.6m word tokens.

„Nacional“ 1997-2000, approx. 6.9m word tokens.

„Fokus“, 2003-2005, approx. 2.7m word tokens.

#### *Other*

- A quarterly magazine published by the influential Croatian cultural institution Matica Hrvatska containing in-deep articles on cultural and societal issues: “Hrvatska revija”, 2001-2005, approx. 1.4m word tokens.

- The Official Gazette of the Republic of Croatia „Narodne novine“, 1990-2005, approx. 18m word tokens.

- Canonical Croatian literary production „Croatian classics“, 1556-1950, approx. 3.6m word tokens.

Three of the four functional styles, i.e.: *Literary*, *Administrative* and *Publicistic* are well represented, whereas the *Scientific* style, which one would find in scientific periodicals, is only represented by “Hrvatska Revija”.

#### *The sources in SrpKor*

The version of the SrpKor used in this study is from 2013 and it consists of 4889 texts. 4534 texts (92.74%) were originally written in Serbian. 355 texts (7.26%) are translations into Serbian. The exact number of *word tokens* is 122,255,064 which represent 1,402,664 *word types*. The complete number of tokens or types, including punctuation, hyphens, etc. the number of tokens is 152,540,721 and of types 1,424.899. (Utvić 2014: 245–47)

SrpKor 2013 is compiled of texts in the *Literary style* written or translated into Serbian in the 20<sup>th</sup> and 21<sup>st</sup> centuries (7.12%), scientific and popular-scientific texts in the *Scientific style* from different domains (natural and social sciences) (3.87%), texts in the *Administrative style* (18.88%), and finally what the corpus engineers call “general texts”, which is an umbrella term used to cover texts from printed and electronic news media published between 1991 and 2012, in other words texts in the *Publicistic style*, (67.37 %). The remaining 3.76% are not assigned any of the functional styles. (Ibid) 120 (2.45%) of the 4889 texts were published before 1990<sup>74</sup>, 3644 (74.54%) were published 1990-2013<sup>75</sup>. 1125 (23.01%) of the texts are not dated. (Ibid: 248)

---

<sup>74</sup> 8 (0.16%) before 1950 and 112 (2.29%) between 1950 and 1989

<sup>75</sup> 1990-1999: 3.19%, 2000-2009: 55.31%, 2010-2013: 16.04%

The factional and some of the fictional texts are, as are the factional texts in the HNK 2.5, collected from newspapers and magazines. The publications, in the following list, which do not have a period assigned, were all published between 1991 and 2002.

*Daily news*

News agency: "TANJUG", 1995-1996

Newspaper "Politika", 2001-2003 and 2005-10

Newspaper "Večernje novosti", 2008-2011, "Danas", 2002-2006

Web-based news portal "Peščanik", 2012

*Weekly magazines*

"Ekonomist" (economy) and its inset "Ebit" (economy and the Internet)

"NIN" (socio-political news)

"Ilustrovana politika" (societal news)

"Moje srce" (romance)

"Svet" (glossy magazine)

*Biweekly, monthly and quarterly magazines*

The official Gazette of the Serbian Orthodox Church "Pravoslavlje"

"Republika" (political, cultural and societal)

"Kalibar" (guns and hunting), "Viva" (health)

"Mostovi" – (scientific magazine on foreign literature and translational issues)

"Teološki pogledi" (scientific, theological journal)

*Other*

The Official Gazette of the Parliament of Serbia "Zakoni Narodne Skupštine Republike Srbije"

"Danica" (popular scientific yearbook on literature, culture and science)

As in the HNK 2.5, the number of word tokens per source is not available, but Utvić (2014) provides statistics on the distribution of texts, word tokens and word types according to the functional styles as cited above.



### Annotation

Both corpora have been lemmatised and the word tokens have been ascribed word category (part-of-speech) and are available for querying using the corpus querying system IMS OCWB<sup>76</sup> The lemmatisation and tagging, i.e. the grouping together of different inflected forms of a lexeme so they can be analysed as a single item and the subsequent or concurrent description of these wordforms, has in both corpora been carried out with a stochastic (probabilistic) tagging programme. SrpKor has been tagged using the programme *TreeTagger* and HNK has been tagged using the programme *CroTag*. *CroTag* was along with The Croatian Morphological Lexicon constructed within the HNK-project. (Tadić 2009: 222–23) Similarly, *TreeTagger* has been adapted to Serbian morphology and has been used in combination with The Serbian Morphological Dictionary to tag a smaller Corpus (*The Serbian Lemmatized and PoS Annotated Corpus*)(<sup>55</sup>) The two morphological dictionaries/lexicons comprise of computational models from Croatian and Serbian inflection, which makes it possible to search the corpora not only for any part of speech or any lemma, but also, for instance all nouns which have the following attributes: *type*: common or proper, *gender*, masculine, feminine or neuter, *number*: singular or plural, *case*: nominative, accusative, genitive, dative, instrumental, locative or vocative and finally *animacy*: animate or inanimate. Depending on the word category different attributes are ascribed. However, the version of the SrpKor used in this investigation has not been tagged in this detailed manner partly because tagging with a stochastic tagging programme entails a large risk of less accuracy in the tagging (Utvić 2011: 42a–43a). The detailed automated tagging of the HNK has an accuracy of this automated tagging which lies between 86 and 98%, being most accurate in tagging of adjectives, nouns and pronouns (Tadić 2009: 222–23).

Even though SrpKor 2013 is without the detailed automated tagging, all tokens have been tagged according to a basic tag-set, which includes 16 tags, 12 of which corresponds to the basic tag-set of the HNK. In other words, it is possible to conduct parallel searches of HNK v.2.5 and SrpKor2013 using the following attributes: *noun, adjective, verb, pronoun, numeral, preposition, conjunction, interjection, particle, adverb, abbreviation* and *unknown*. <sup>77</sup>Utvić 2011: 42a–43a)

---

<sup>76</sup> IMS OCWB is an abbreviation of: Institut für Maschinelle Sprachverarbeitung Open Corpus Workbench. A tool developed at The University of Stuttgart.

<sup>77</sup> The additional basic tags available when querying SrpKor are: *punctuation, Roman numeral, prefix* and *sentence end marker*.

Every text in SrpKor is also bibliographically annotated, so that one may inquire for its title, its year of publication, its author, the functional style, and whether it was originally written in Serbian. In HNK every text is annotated with abbreviations from which the source may be deduced, e.g. *nn1997* which refers to “Narodne Novine” published 1997.<sup>78</sup> Similarly, the individual texts in the HNK 2.5 are not assigned any of the functional styles, but we know that all four written functional styles are present (cf. p. 48) and I will therefore be able to deduce the functional style in individual texts, when necessary, on the basis of the bibliographic information.

In the following tables (Table 1.1-1.3) I have summed up the information about the corpora and information extractable from the two corpora.

---

<sup>78</sup> <http://hnk.ffzg.hr/struktura.html> (Accessed July 13, 2015)

Table 1.1 - Quantity

	<i>SrpKor 2013</i>	<i>HNK 2.5</i>
tokens	152,540,721	not specified
types	1,424,899	not specified
word tokens	122,225,064	100,884,824
word types	1,402,664	1,153,226
no of lemmas	not specified	not specified
no of texts	4889	not specified
original texts	4534	all
translated texts	355	none
texts publ. prior to 1990	2.45%	not specified
word tokens publ. prior to 1990	not specified	3.67%
texts publ. after 1989	74.54%	not specified
word tokens publ. after 1989	not specified	96.33%
no date	23.01%	none

Table 1.2 - Quality

	<i>SrpKor 2013</i>	<i>HNK 2.5</i>
PoS-tagged	yes	yes
lemmatised <sup>79</sup>	yes	yes
bibliographic info	explicit	deducible
functional style	explicit	deducible

In the following table you may see how comparable the two corpora are with a view to quantity and representativeness of the four written functional styles.

---

<sup>79</sup> You may search for all forms of a given lemma, for instance all forms of the noun "hotel"

Table 1.3 - Comparability

Functional style	Corpus	Word tokens	
Literary	SrpKor 2013	13,9m	11.34%
	HNK 2.5	3,6m	3.57%
Scientific	SrpKor 2013	4,2m	3.45%
	HNK 2.5	1,4m	1.39%
Publicistic	SrpKor 2013	90,1m	73.68%
	HNK 2.5	77,8m	77.18%
Administrative	SrpKor 2013	6,9m	3.33%
	HNK 2.5	18,0m	17.86%
Other	SrpKor 2013	7,2m	5.90%
	HNK 2.5	0,0m	0.00%

We find the biggest discrepancies in sources written in Literary and Administrative style but the major part of both corpora, which is also the part in which most linguistic innovations are expected to surface, is the publicistic one (SrpKor 73.63% and HNK 77.18%).

From “Table 1.1 – Quantity” it is also very clear that both corpora primarily consist of text or tokens in text published after 1989 (HNK 96.33%, SrpKor 74.54%)<sup>80</sup>

In “The Handbook of Language Variation and Change” (Chambers, Trudgill, and Schilling 2002), Laurie Bauer in his contribution, “Inferring Variation and Change from Public Corpora” defines the public corpora as “a body of data which can serve as the basis for linguistic analysis and description and which is available to linguists in general as an identifiable whole or from easy accessible materials” (Bauer 2002: 99) Bauer, then, proceeds to categorise different types of corpus. The HNK and the SrpKor are according to his classification *simple, electronic, textual, structured, written language, public corpora* (cf. Bauer 2002: 100-102).

HNK and SrpKor are *simple* as opposed to *comparative* by being corpora that are representative of the language variety in a rather short period of time and primarily representative of one functional style. In other words they are not suitable for researching varieties on a diachronic axis, and only to a limited extent suited for researching variation between styles. Both corpora are *structured* with a view to the texts themselves and in that they are annotated as described above. And finally, they are *textual* in that they comprise texts not word lists and these texts are an expression of *written language* except in very few instances,

---

<sup>80</sup> The actual percentage of texts in SrpKor 2013 published after 1989 may be as large as 97.25% (thus equalling the percentage in HNK 2.5), as 23.01% have not been assigned a year of publication.

where verbal dialogue or monologue is directly quoted in, for instance, a newspaper article, and even then the utterances have been adapted to the written language and do not reflect the phonetic features of the utterances, which would be the case in a transcribed *spoken language* corpus.

Bauer then informs us of what the problems as well as benefits of a public corpus may be. (Ibid: 102-104). For one, it is possibly misleading to conclude anything on the basis of number of samples alone. "The researcher needs to ask how far the corpus reflects anything but the collection of texts/words that make the corpus." (Ibid: 103) I believe that this particular problem is and will be solved in this study by the description of the texts in the corpora (above) and by providing additional description of the source texts of any less attested, but for the study significant findings. An example of such a situation is the 6 (or 9.8 pct.) occurrences of the replicated noun *bookmaker* (< bookmaker)<sup>81</sup> in the HNK, and the more expected 55 (90.2 pct.) occurrences of the domestic substitute *kladioničar* ('bookmaker, betting agent'). In order to ensure that these six occurrences are not either from an English quote, from the same text, or from the same publisher (i.e. newspaper), all six occurrences were analysed in their broader context. I could conclude that four of the occurrences were in the same daily newspaper (*Vjesnik*), but in different issues (11 Dec, 2002, 18 Aug 2002, 20 Jun 2002, 29 Nov 2000) whereas the last two occurrences were from other publications (*Nacional* and *Glas Slavonije*). Furthermore two of the occurrences were put in quotation marks which I deduce is the author's way of indicating that he/she is using a foreign word, a technique which is in line with the recommendations of the linguistic authorities in Croatia (cf. subsection 4.2.1, p. 159). Lastly, I could conclude that one of the six occurrences was spelled phonologically, as is recommended in the Serbian standard and regarded as incorrect in the Croatian standard. The conclusion for this particular pair is that the replicated word *bookmaker* is in very scarce use as opposed to its domestic counterpart and, when it is used in the Publicistic functional style some authors show a high awareness of *bookmaker's* foreign origin and therefore adhere to the standard language norm of indicating this fact by orthographic means. It follows that the Croatian usage norm represented in the HNK concerning the use of the pair *bookmaker/kladioničar* strongly prefers the domestic *kladioničar* and the occasional use

---

<sup>81</sup> In the meaning: a person whose job is to take bets, calculate odds, and pay out winnings.

of the replicated *bookmaker* should be seen as deviating from the prevailing usage norm. It could then be speculated that the occasional uses of *bookmaker* are indicators of the spoken usage norm, but as the corpus is only representative of the written language it will, in this study, remain a speculative claim. In the SrpKor *bookmaker/bukmejker* is in much larger use. 33 occurrences were attested there.<sup>82</sup>

There are also possible problems when comparing two (or more) corpora. The different text selections in the searched corpora may be the underlying reason for the different findings, rather than the difference or similarity in the usage norm of the languages. (Bauer 2002: 103) Again, the structural likeness of the two corpora in the present study, as shown in Tables 1.1.-1.3 above, should prevent the actuality of this problem, and when specific results could be decisive for the overall conclusion considering a searched phenomenon, those results will be scrutinized in a way similar to what I have just described in order to establish whether the results are limited to one author, one publication, or one functional style in either of the corpora.

A third potential problem is the size of the corpora. "Corpora used to investigate lexical matters generally have to be extremely large [...]; some grammatical phenomena are also so rare in texts as to require very large corpora if reasonable amounts of data are to be found" (Bauer 2002: 104). I believe to have avoided this potential problem by having accessed two very large corpora, and by having designed my searches and my analyses of the results in a manner that excludes results that are based on too small an amount of data. I have in Bauer's words dealt with the statistics carefully, so they need not prevent similarities and differences from being discovered (Ibid.).

The benefits of public corpora are that all my searches may be replicated by other scholars and that I may treat my phenomena numerically (Ibid: 102-103). To gain access to HNK, one simply needs to go on-line and type in the address: <http://www.hnk.ffzg.hr/>. To access SrpKor, one needs to procure a password which is readily given to any serious scholar. The webpage through which you may contact the corpus administrators is: <http://korpus.matf.bg.ac.rs/>.

In conclusion, I hold that the electronic text corpora presented here: the HNK and the SrpKor, despite their deficiencies, meet the demand to be

---

<sup>82</sup> For technical details on how I queried the corpora for the investigated phenomena, please cf. Appendix D, p.333ff

representative of a period between 1990 and 2005 in addition to containing earlier and later text. They are of a size which is sufficient to provide indicative results in investigations of lexical and grammatical phenomena. They are both mostly representative of the language of the media, i.e. the Publicistic functional styles of Serbian and Croatian. Due to their annotation, they are searchable according to identical search criteria. It is therefore possible to compare the results of the corpora in a scientifically valid manner, as long as the researcher does not treat the results in a purely numerical/quantitative manner but adds a qualitative analysis of the findings whenever needed.

The details concerning the different types of searches, that I have conducted, will be explained in connection with descriptions and analyses of the findings regarding each of the types of linguistic innovations, which ensue in chapter 4.<sup>83</sup>

After having introduced the history and variational differences of Serbian and Croatian and the empirical data as well as the linguistic phenomena under investigation on the previous pages, we will now turn to the field of language planning and language policy which will provide an insights into the extra-linguistic measures taken to establish and maintain the form and status of a particular language while also discussing the actors that initiate and implement these measures including reasons and motivation which drove, have driven and are still driving the actors in language planning.

---

<sup>83</sup> Specifically in the following four subsections about the inspected empirical data: 4.3.1, 4.4.3, 4.5.4 and 4.6.3.

## 2. Language Planning and the Standard Language

In our ever-shrinking world the notion that a national language is something worth working at and struggling for may seem merely a piece of inscrutable mysticism.

(Haugen 1966: 2)

In this chapter I will, after in section 2.1 having discussed different definitions and uses of the term standard language – the object of language planning, in section 2.2 attempt to shed some light on the concept of language planning first by discussing the purpose(s) of language planning by putting forward some of the views held by scholars in the field of language planning and language policy (LPLP). Second, in section 2.3, I will look into the categorisations within language planning (LP) and discuss issues concerning status, structure and function of languages and introduce concepts such as *corpus planning*, *status planning*, *acquisition planning* and *prestige planning*. Thirdly, in section 2.4, I will go through the phases or steps of which language planning consists, both with a view to initial standardisation of a language but also, equally important, with regards to the later changes within a standard language.

I will use a framework proposed by a leading language planning expert from the language area, the sociolinguist Milorad Radovanović, professor at University of Novi Sad, Serbia along with suggested explanatory frameworks put forward and referred to by Kaplan and Baldauf in “Language Planning – from theory to practice” (1997), by Haarmann (1990) and finally by Hornberger (2006) in her attempt in “Frameworks and Models in Language Policy and Planning” at integrating the views and models of a range of LP-theoreticians. Of special interest will be the phases of *elaboration* or *cultivation* of a language and of *implementation* as it is within these phases of adding and changing a language and the subsequent usage of these additions and changes, that we find the phenomena which, in a more concrete form, will be the focus of attention in my investigation of contact-induced language changes in Serbian and Croatian. The theoretical frameworks used by LP-theoreticians will serve to illustrate which are the forces that may or may not motivate the addition, recognition and subsequent acceptance of a linguistic innovation within a standard language. I will, thus, by way of the different frame-



works, illustrate how the national standard languages Serbian and Croatian were, have been and indeed still are formed through this process. I will take the reader through the phases which a language variety usually goes in order to be promoted to the function and status of a *standard language*. The crucial part of the planners, the *agents* of planning will be illustrated and exemplified by analysing the Croatian and Serbian agents of language planning and policy. In other words: How the Serbian and Croatian standard languages come about and who decides what proper Serbian is and what is proper Croatian. The desired outcome of language planning should, as mentioned, be a standardised form of the language in question – a standard language.

The terms *standard language*, *standardisation*, *standard variety*, and *norm* are frequently used in the Language Planning literature which is why a clarification and discussion of the standard language ensues. In section 2.1.1 “System, norm, speech”, before I return to issues of language planning itself. The crucial concept in this discussion is the concept of norm, so when defining and planning (for) a standard language, the question of what a norm is, and what the norm of the language should be, naturally arises. To answer this question I will apply Eugenio Coseriu’s view of linguistic norm and discuss it vis-à-vis the elaboration of Coseriu’s concept of norm, done by the Croatian sociolinguist Josip Silić. (Coseriu 1974; Silić 2006)

Generally speaking, language planning is a phenomenon studied in sociolinguistics, which is the branch of linguistic research which focuses on the speech community. A speech community may be as small as a village or a club or as large as a nation or even a group of nations (Southerland and Katamba 1997: 540).

Language planning is, in its simplest sense, an attempt by someone to modify the linguistic behaviour of some community for some reason (Goebel et al. 1997, 2: 3). And language planning in its most straight-forward form begins with the standardisation of an idiom, thus promoting or converting it into a standard language. The standardisation process may take place over a long time and not necessarily ever lead to an officially recognised standard language.

Early studies of language planning focus on the “establishing” of a standard (national) language and are therefore mostly a thing of the past, especially in Europe, whose peoples began identifying as nations/nation-

alities based on ethnicity already in 18<sup>th</sup> and 19<sup>th</sup> centuries. This is however not the case with the standard languages Serbian, Croatian, Bosnian and Montenegrin which have (re-)gained their status as standard languages during the last two and a half decades, cf. p. 9.

Furthermore, language planning is the continuous development and maintenance of a language which is often happening concurrently and may even be considered a precondition for gaining the status of standard language. So, the planning of Serbian and Croatian is not only an activity of Serbian's and Croatian's past, as shown in section 1.1 above, but also an on-going activity which aims to modify the behaviour of a speech community.

Language planning is therefore closely linked to language policy. Firstly, the policy-makers concern themselves with inaugurating variety as the official language, secondly through legislation and political as well as financial support they influence the way in which the corpus of the language is planned.

Initially the term 'language planning' referred only to government and/or national 'top-down' planning but it is now also used to refer to many other ways of looking at the planning of language both on the micro-level and as 'bottom-up'-planning (Kaplan and Baldauf 1997: 27). As we are dealing with planning of the national standard languages, Serbian and Croatian, the main focus will be on this so-called 'top-down' planning and the possible influence of the common speaker/speech community on language planning will be viewed as a reaction to or reflection of 'top-down'-planning.

## **2.1 Standard language and norm**

In the language planning efforts in the 19<sup>th</sup> century, the object was to create a common language norm for all the members of a given speech community, so that they could truly become members of the same nation and therefore rally around the goal of becoming a state. However, at the time the term standard language had not yet been coined. The normative works among Croats and Serbs alike referred to the literary language – *književni jezik* instead. The languages Serbian and Croatian have, as does the majority of what are called languages, a *standard variety*, nowadays commonly known as the *standard language*. A standard language is then the product of standardisation. The term *standard language* is a term first used in American sociolinguistics (Kristiansen 2001: 348). And, according to Haugen (1966), it holds true for all standard languages that

their planning is not only a linguistic but also pre-eminently a socio-political problem (Haugen 1966: 3)

*Standardisation* is also a frequently used term and it may, as we will return to later, be viewed as representing the part of the activities subsumed in language planning pertaining to a) obtaining the official status of a standard language and b) the maintenance of an already existing (official and/or national) standard language.

Stewart (1968), quoted by Fishman (1975), defines standardisation not only as the codification" but also as "the *acceptance*, within a community of users, of a *formal set of norms defining 'correct' usage*" (my emphasis) (Fishman 1975: 24). Stewart's *correct usage* corresponds to Hornberger's *uniformity* when she claims that standardisation of the language refers to "language-planning activities that codify the linguistic forms of [a] standard as a uniform norm.", and she thereby omits the acceptance by language users. (Hornberger 2006: 31). Milroy (2001) shows us that uniformity might not be the primary aim of standardisation, when he argues that linguists use the terms *standard*, *standard variety* and *standardisation* very differently, sometimes contradictory because a standard variety can be taken to mean a) a *uniform* norm with the least possible variation and b) the most *prestigious* variety. A prestigious language variety is not necessarily the most uniform language variety as the variety is prestigious not in itself but because of the prestige of the users of this particular language variety. (Milroy 2001: 531). However, when standardisation is defined as the process leading to a standard whether it is standard for a language or for how to build a house, standardisation imposes uniformity, and thereby limits variation. John E. Joseph puts forward a similar point as the process of standardisation, in his words, is "an intensely political process", in which one dialect/variety is selected, thus oppressing all others, after which variations within the selected language variety are also oppressed so that uniformity can be achieved (Joseph 2006: 9).

In the standardisation of the Serbian and Croatian languages, and I suspect in many other languages, both ideals were in play, uniformity *and* prestige.

The standardisation process may, as mentioned above, be viewed as the process aiming at establishing a standard language or/and as the process in which the norm of an existing standard language is defined. In order to distinguish between the two processes, Peti-Stantić (2008) pro-

poses the terms *primary* and *secondary standardisation*. Primary standardisation is a political project which has a unifying function, i.e. to unify people within one nation by means of a language, that this nation can call its own. The secondary standardisation has as its aim to fortify an already existing standard language by emphasising its uniqueness vis-à-vis other similar or related standard languages. Whereas the primary standardisation has unification as its goal, secondary standardisation often has the opposite purpose – diversification. (Peti-Stantić 2008: 72–73). The purpose of secondary standardisation is, in other words, to build up a language in opposition to similar language varieties which is what Heinz Kloss in 1967 termed *Ausbau* and speculated that without *Ausbau*<sup>84</sup>, that is, standardisation undertaken in different geographical areas with similar languages (for instance Sweden and Denmark, Iceland and the Faroe Islands or in the Czech and Slovak areas) there would only have been three and not six structurally different languages today (Kloss 1967: 30).

One might see the processes of primary and secondary standardisation as opposites, as Peti-Stantić does, but I hold, that it is equally true that both primary and secondary standardisation have unification as their goal. The difference lies, rather, in what is (or who are) being unified. In the Serbian, Croatian case, primary standardisation was aimed at unifying Serbs and Croats, whereas the secondary standardisation besides being aimed at diversification from the Serbs, also aimed at uniting the Croats. In both cases uniformity was imposed, although on different levels. At the early stage it was a question of selecting a unifying language variety and a unifying orthography. Attempts at further unification were less successful in the long run, as we saw in section 1.1 above. At the later stage, unification is sought by stigmatising what is considered improper and simultaneously offering “proper” and uniting Croatian alternatives.

Some scholars prefer the term *standard variety* (of a language) as opposed to *standard language*, thus emphasising that a language consists of many varieties of which one is the standard. The standard variety is the variety of a language which has first of all been selected to fill this

---

<sup>84</sup> The distinction between *Ausbau* and *Abstand*-languages will be explained in section 2.3.

role. Other steps ensue in order for the standard variety/language to fulfil its functions<sup>85</sup>. One could say that language planners through *standardisation* of a language variety aim at creating a *standard language*, i.e. promoting a language variety to the status of *standard variety*.

Depending on whether a) you look upon the process of developing a standard language, i.e. primary standardisation/standardisation of status or b) you look upon the standard language as a given as opposed to other varieties of a given language/standardisation of corpus or c) you are more interested in the functions a standard language has to fulfil - the defining features are quite diverse.

*a) A standard language as the result of language planning*

Haugen (1966) and many others following him explain how a standard language emerges as a result of language planning.

Language planning can be described in steps beginning with the selection or modification of an existing variety (in Haugen's words: *selection of norm*) which is then implemented through the *codification of form* which refers to formal codification of script, orthography, pronunciation, grammatical forms and lexical items (Haugen 1966: 18–21). In order for a language variety to function as a standard language it also needs to be *elaborated*, i.e. constantly be adapted to the communicative needs of its speakers by producing new linguistic matter to denote new concepts. This labelling of new concepts can come about in different ways. If a concept already exists in another variety or language the foreign wording may simply be adapted (*adaptation*) or a new word will be invented (*innovation*).<sup>86</sup> However, the selection, codification and elaboration of a language variety do not make a standard language. The fourth key element of language planning is (in Haugen's words) *Acceptance by the community*. In other words, if the planned language is not used by the speech community, it will not be a complete standard language.

*b) A standard language as opposed to other varieties*

Some scholars choose to focus on the relation between the standard and non-standard varieties of a language and tell us that standard language

---

<sup>85</sup> These steps will be further illustrated in section 2.4

<sup>86</sup> According to my definition, both phenomena, i.e. replication with ensuing adaptation of lexical and grammatical from another language or language variety as well as the coining of or inventing new lexical and grammatical items with language-internal means, constitute linguistic innovations

is the language variety that ranks above other varieties. A *standard language* may be defined as an entire language which has among its varieties at least one *standard variety*. [...] [O]n higher levels of cultivation [...], standard varieties are virtually always used," (Ammon 1989: 86)

"This superposed variety is employed by the government and communications media, used and taught in educational institutions, and is the main or only written language." (Southerland and Katamba 1997: 541).

The standard "designate[s] a *variety* of a language (which follows a 'norm' or 'codex', i.e. 'standard' does not designate the norm itself)" (Auer 2005: 8). Speakers of different varieties orientate to the standard variety, which is viewed as prestigious and it is/has been codified and elaborated (Auer 2005 *ibid*)

As the phenomena under investigation in this thesis primarily have to do with the relation between standard languages and linguistic innovations induced by contact with foreign languages, the relation to and possible influence induced by non-standard varieties of the standard language in question are not taken into account. Only if we were consider Serbian a variety of Croatian or vice versa (which some less acknowledged Serbian scholars do) could we consider applying this kind of distinctions.

### *c) A standard language according to its functions*

Other scholars choose to focus on functionality and describe standard languages as language varieties that fulfil certain functions in society. Among them we find several Serbo-Croatian and Croatian linguists. Silić (2006), for one, regards a standard language as a polyfunctional institution and as such its function is to be a means of communication between human beings in all areas of life. (Silić 2006: 10) Similarly, Mićanović (2006) states that a standard language stands out because of its functional role and the communicative demands it must fulfil. (Mićanović 2006: 559). According to Brozović (1970), a standard language cannot be defined in the same way as idioms are defined in dialectology (historical linguistics) by delimiting dialects and categorising languages in genetic language families or in comparative structural studies (linguistic typology), where languages are categorised in types according to their structure and substance. Neither the linguistic structure of an idiom nor its genetic closeness to others need play a role when a standard language takes it form. In contrast, a standard language is the product of extralinguistic factors. In other words the "standardness" of an idiom has

nothing to do with linguistic factors (i.e. its structure and substance). It has to do with use and function instead. If and only if a language variety is used as a means of communications within a civilisation (an ethnic or national formation) and if it has the required functional qualities to be a means of communication, it qualifies as a standard language.

There are three key factors at play when assessing whether a given language variety can be regarded as a standard language. In order for a language variety to be regarded as a standard language it has to a) be *the communicative and expressive means of a civilisation*, b) have *the functional qualities* which enable it to be the *communicative and expressive means of a civilisation* and c) have the character of the civilisation itself (Brozović 1970: 28). So, a standard language, much in line with Kloss' definition of Ausbau languages, is defined by being autonomous, having a norm and being polyfunctional. By claiming that the functional and formal aspects of a language variety are what define it as a standard language, Brozović ignores the linguistic criteria applied when determining whether a language is a language in itself or a variety of a language. The general criterion is "the proportion of sameness and difference" (Kloss 1976: 303) covering some more specific criteria such as "the dialectal basis" (Kloss 1976: 310), "interintelligibility" (Wayles-Browne, quoted by (Kordić 2010: 102), „syllable structure" (Jakobsen 2008). According to these criteria Serbian and Croatian *are* one language, as the two standard languages have the same dialectal basis, speakers of the standard variety of Serbian and Croatian understand each other effortlessly and structurally the syllable structure of Serbian and Croatian is the same. The tendency, especially within the Croatian linguistic normative bodies, is, not surprisingly, to promote the usage of words, affixes, declensions and syntactical structures which appear archaic, artificial or foreign to Serbian usage norm. This tendency seems to validate the need for structural differences. In other words, it is not sufficient for the normative bodies or the average Croatian speaker to proclaim that Croatian is a separate standard language which fulfils all the necessary communicative functions. Contrary to Brozović's claim, structural differences are necessary to validate this claim.

Thus, the notion of standard language can be approached in different ways. However, the authors seem to be in agreement that the standard is superposed and is meant, functionally, to overarch (geographically, socially and/or stylistically) other varieties/versions of the same language. It may prove difficult to delineate the standard language (variety)

from other forms (varieties) of the same language. Usually, the standard language encompasses what is deemed correct or proper language use in dictionaries, grammars and orthographic manuals (a prescribed norm) and by organisations and individuals who are regarded as authorities on linguistic matters (a model-based norm). Silić (2006) employs and develops ideas from Eugenio Coseriu's theory on language and language change in order to reach a definition of the standard language, which will be discussed in the following section.

### 2.1.1 System, norm, speech

Eugenio Coseriu has in his theory on language and language change a very illustrative tripartite definition of language. Language may thus be viewed in the trichotomy: system|norm|speech. Coseriu aptly inserts the category *norm* in between the Saussurean categories for *langue* and *parole* and thus makes us aware that there are regularities influencing the langue/system which do not stem from the parole/speech, but rather from the extra-linguistic communicative and symbolic needs and habits of the speech community.<sup>87</sup>

The norm is the realisation of what is socially and culturally established as opposed to the language system, which is the term used to denote the inherent rules and regularities (techniques) which in turn delimits what is possible to do morphologically (in word-formation, declension and conjugation), syntactically (how words and phrases are combined), phonologically (which sounds are used to distinguish between different meanings), phonetically (which sounds are utterable) within a given language system. The norm is what (according to these rules and techniques) has already been put into a realised pattern. The system represents the dynamic (the possible outcomes) whereas the norm represents the fixation (the acceptable outcomes). Speech is the concrete manifestation of language which may violate both what is acceptable according to the norm, and what is possible according to the system. (Coseriu 1974: 47–58).

The norm varies along three axes, the diatopic (geographically), the diastratic (socially) and the diaphasic (stylistic). This entails that not

---

<sup>87</sup> Coseriu's definition of norm has also been applied in lexicology when accounting for the arbitrariness in the usage of synonymous derivational affixes, and the unacceptability of applying the "wrong" affix. The norm accounts for the choice between alternative word-formation types (to nationalize, to clean, but not \*to nationality, \*to national) (Lipka 1992: 96)



only standard languages are used according to a norm, but all varieties, be they dialectal, sociolectal or stylistic are used according to norms, which are violated if you utter something unacceptable. Coseriu clarifies his notion of norm defining it as *die exemplarische Sprache*. *Die exemplarische Sprache* of a speech community is not realized in actual speech but is solely the *norm* towards which speakers orient their speech ('Sprachliche Varietäten' 2015).

The belief that the term *norm* should cover what is accepted in actual language use and not what is prescribed to be correct in normative works is also shared by the Prague School (Neustupný 1989: 213). Bohuslav Havránek, a member of The Prague School<sup>88</sup>, tells us, as early as 1936 that every speech community (Sprachgemeinschaft), regardless of its size and its location and whether or not it has a writing system, has a language norm (Sprachnorm) and that every member of a community will accommodate his or her usage to the norm of the community. Havránek adds that it is irrelevant whether this phenomenon is called a standard language (Standard) or a language norm (Sprachnorm) (Havránek 1964 [1936]: 414). According to another prominent member of the Prague School, Vilém Mathesius, in order for the language norm, or indeed, the standard language to function at optimum level, it needs to be both stable and flexible. Mathesius' term (1932) – 'flexible stability' refers to the notion that a standard language has to be flexible so that it may adapt to the constant change in communicative needs of its users. And at the same time the standard language has to be stable so that its users (the speech community) do not lose touch with their past due to an inability to read older texts in the same 'language'. (Gammelgaard 2004: 66)

If a standard language is a reflection of Coseriu's norm, it follows that it cannot be concrete. As soon as the norm is put into concrete utterances (written or verbal) it becomes speech. In the strictest interpretation (being that standard is equal to norm) an utterance or text cannot be produced in the standard language. However, in grammars, orthographies and other manuals on language, it is frequently stated or under-

---

<sup>88</sup> A school of linguistic thought and analysis established in Prague in the 1920s by Vilém Mathesius. It included among its most prominent members the Russian linguist Nikolay Trubetskoy and the Russian-born American linguist Roman Jakobson; the school was most active during the 1920s and '30s (Encyclopaedia Britannica 2015).

stood that what is described is the standard language. Thus, the standard language is expressed through speech, but only the speech that is produced according to the rules of the system and the norm (the acceptable). Accordingly, the term standard language and norm are in Serbia and Croatia often used to refer to the sum of the rules and “acceptable outcomes” (e.g. (Radovanović 2009; Silić 2001; Matasović 2011). This understanding of the term *norm* is generally also present in literature on language planning. Wright (2004: 54) says: “The norm is decided and codified by a central group, disseminated by the institutions of the state such as education and the usage is constantly policed and users dissuaded from divergent practices, both formally and informally.” Again, this is not a new perspective. In 1936 Havránek pointed out that, unfortunately, the normative works (grammar books, etc.) frequently are seen as the norm itself (Havránek 1964 [1936]: 414)<sup>89</sup> as does for instance Matasović when he says that ‘standard languages are defined by normative grammars and dictionaries’<sup>90</sup> (Matasović 2011: 84). Havránek believes that the norm should be viewed as a phenomenon which arises within a speech community without a particular agent, motivated by the needs of a speech community, alternatively as a norm set up by experts in language who interprets the needs of a speech community. The needs of the speech community will in both instances be communicative as well as symbolic (cf. section 2.2 below)

Josip Silić (2006) combines the oppositional views of what defines a standard language and thus uses the term ‘standard language’ as an umbrella term for both interpretations of norm. He develops Coseriu’s trichotomy to account for this and defines three subcategories within standard language:

*Figure 2.1 - Standard language according to Silić*

system: codification : norm : usage : speech.  
standard language

---

<sup>89</sup> “Eine solche Kodifikation, gewöhnlich in Schulgrammatiken oder anderen Kompendien, hält man oft für die Sprachnorm selbst, aber danach würde die Sprachnorm eine aussersprachliche Erscheinung darstellen.”

<sup>90</sup> Original: “[Standardni jezici] su u pravilu definirani svojim normativnim gramatikama i rječnicima.”

By inserting the terms 'codification' and 'usage', Silić thus narrows the scope of Coseriu's norm. Silić also views the terms as covering different levels of abstraction:

*Figure 2.2 - Language on five levels of abstraction*

Speech – Concrete  
Usage – 1<sup>st</sup> level of abstraction  
Norm – 2<sup>nd</sup> level of abstraction  
Codification – 3<sup>rd</sup> level of abstraction  
System – 4<sup>th</sup> level of abstraction

As can be seen from the illustration the standard language encompasses the 1<sup>st</sup>, 2<sup>nd</sup> and 3<sup>rd</sup> level of abstraction.<sup>91</sup> Silić's different levels of abstraction should, in my view, despite of the numbering not be considered as a cline, which a linguistic element has to climb in order to become part of the *system*, but rather as a useful specification of different status of an element within the standard language, i.e. 1<sup>st</sup> to 3<sup>rd</sup> level of abstraction.

I will, in my investigation of contact-induced language changes in Serbian and Croatian, via empirical data attempt to ascertain how the norm is manifested in usage and, when appropriate, compare it to the codification. In other words, my findings in concrete speech will illustrate changes in the usage of the standard compare this usage to the codified norm even allude to the highest level of abstraction, the *system* and thus be able to reach conclusions about innovations in the systems, i.e. the structure of Croatian and Serbian, independently of whether these innovations have entered the codified norm.

Standard Croatian and standard Serbian will, in other words, in this thesis not be interpreted only as the described and prescribed norm (codification), but as the norm which manifests itself in the language usage in the electronic text corpora. I chose to call this kind of language norm the *usage norm*. The observations will, as mentioned in the introduction, be done by searching through electronic large text corpora of Serbian and Croatian for specific changes. The results and analysis of the searches will help us determine whether changes in the usage norm have occurred, how (and if) these changes contribute to the diversification or

---

<sup>91</sup> Silić alters Coseriu's definition of norm, so that may fit into his framework. Coseriu's definition of norm, seems to cover all that is acceptable to the language users, not only what is acceptable in the eyes of the language codifiers.

unification of Serbian and Croatian standard languages, whether prescriptive codifiers (normativists) and language planners have an impact on usage of contact-induced changes and finally whether the linguistic system (the grammar) of Serbian and/or Croatian are changing.

Now that we have an applicable distinction between usage norm and prescribed norm, let us take a closer look at the purposes of language planning.

## 2.2 Identity and communication

Whether language is viewed as an innate ability (a biological phenomenon) or as an ability which is learnt by humans to be able to communicate (a social construct), it is formed by its users, the speakers (Wright 2004: 5). How the speakers form their language is dependent on what they need to communicate. The needs of language users are what translate into the functions needed in a language in order for it to be an adequate means of communication. In this view, to which I adhere, the ideal goal of language planning is to a) ensure a continuous elaboration of the language and at the same time b) ensure that the language eases communication among its speakers. Language is, however, not an activity that one can plan and subsequently carry out. The planning activities can only be an attempt to change and/or develop a practice already existing in some form (Joseph 2006: 9).

Scholars such as for instance Crystal (2008) define the purpose of language planning as “[...] to solve the communication problems of a community [...]” (Crystal 2008: 268–69). However, the choices made in language policy-making may very well reflect another purpose of language planning, i.e. creating a group identity (Wright 2004: 6). Language planning is tightly connected to nation-building. Language planning was in fact “an integral part of nation-building” (Wright 2004: 8).

The Croatian linguist Škiljan<sup>92</sup> (2000) sees the combination of these two purposes, i.e. ease of communication and creation of identity as the real purpose of language planning. Thus, he states that a linguistic community (i.e. speakers of the same language) is situated in two spaces which he (referring to Edwards (1985)) designates the communicative and the symbolic space, because “... the linguistic community often relates to the language, particularly in its standardised form [i.e. the prescribed norm], as well as to the nation.” Many languages with the status

---

<sup>92</sup> Škiljan is also the author of *Jezična politika* ('Language policy') from 1988 and has commented on Radovanović's LP-model on several occasions

of language today have thus, for political reasons, been built up (Ausbau) as such. The motivation for this cultivation can be explained sociologically as Sue Wright does it “Scholarly activists in the many groups seeking to exit from the rule of the British, Ottoman, Russian or Austro-Hungarian and French Empires laboured to codify, standardise and disseminate a single language for the group, which could then be presented as part of the evidence for its claim to a separate polity.” (Wright 2004: 8). Wright is in concord with Einar Haugen who, in his ground-breaking work on language planning: “Language Conflict and Language Planning – The Case of Norwegian” (1966), states that “...[the] problem of language (...) is not purely linguistic. It is also, and perhaps pre-eminently, a socio-political problem, with roots that reach down into the heart of Norwegian life. This is true of standard languages and their growth everywhere, ..”(1966: 3). This view is shared by Czerwiński who in his article “Language Planning and Lexical Networks in Croatian Media” (2004) points to the fact that the product of language planning – the standard language – is not only a social and cultural construct but more importantly a product of choices made at a moment of national revival (Czerwiński 2004: 1).

As these quotes show, planning (including standardisation) of languages, at least in Europe, was very much a part of the nation-building process of the 19<sup>th</sup> century emancipation from the overarching empires. Even though Serbian and Croatian scholars in this period merged their efforts in order to define and emancipate their nations from the Austro-Hungarian and Ottoman Empires, we see a clear tendency to new emancipation in Croatia in the 20<sup>th</sup> and certainly in the 21<sup>st</sup> century, this time from Serbian or unitaristic Yugoslav homogenization of the Serbo-Croat language, but with the same motivation, i.e. emancipation and consolidation of a separate national identity (within or outside of the interwar or post Second World War Yugoslavia) .

Thus, language planning may be a means to emancipation from an “overlord” of a certain polity and simultaneously a means to impose language uniformity among speakers within a new polity, and this polity is in turn defined by nationality. In other words; by planning and thus altering the norm of standard Croatian or Serbian, the mutual understanding (communication) within the speech community ideally<sup>93</sup> increases,

---

<sup>93</sup> The alteration of the norms of standard Croatian, does, however, not always increase mutual understanding within the speech community. (Kordić 2010: 144)

but at the same time the moving away from the standard Serbo-Croatian, the mutual understanding (ease of communication) between Croats and Serbs decreases.

The understanding of language planning referred to here by Haugen, Wright, Czerwiński, Joseph and Škiljan, and also by scholars such as by Ulrich Ammon, Heinz Kloss and James Milroy is language planning with the clear goal of establishing a standardized form of this or that language. This kind of language planning is therefore often referred to as *standardisation*, cf. p. 60. As a consequence, the initial planning of a language variety is closely linked to a desire to give said language variety the status of a language. So, language planning encompasses cultivation (development, elaboration) of a language variety, (primary) standardisation of the language variety (which entails imposition of uniformity and invariance) and its ultimate aim is to provide support to a group identity, e.g. fuelling national revival.

Some scholars, among them Škiljan (2000), claim that any language policy and thereby any language planning cannot be successful unless the symbolic function of said language concerning national or ethnic identity are taken into consideration, which means that if a national language policy is to work the policy makers and language planners must acknowledge that: "The symbolic and emotional dimensions of national identity are crucial, and language policies that ignore them prove dysfunctional in the long run." (Joseph 2006: 24).

So, it seems to be established that the symbolic, identity-related function of a standard (national) language is so influential that it may easily overrule the communicative function and even lead to difficulties in communication, illustrated in Croatia in the late 90s by what Opačić (2002) calls *Strah od materinskoga (hrvatskoga) standardnog jezika* ('Fear of the standard (Croatian) mother-tongue') which resulted in public figures mumbling and/or seeking refuge in their own dialect instead of speaking the standard language when appearing on camera in fear of using a politically stigmatizing word, which could label the speaker as a traitor, communist or worse in the public eye. (Opačić 2002: 229). This counter-communicative politically induced effect on language is puzzling. How come the speakers, the language users, allow their means of communication to be weakened? In the next subsection, the question why people need a norm addressed.

### 2.2.1 Why hamper communication?

If the standardised version, i.e. the prescribed norm, of a language is so rigid that it hampers communication and mutual understanding within the speech community itself, why is it then pursued? Blommaert et al.(2012) attempt to explain it by investigating what it is that makes communities strive to set up a norm, a standard through language planning. They do this not by listing the pragmatic effects (ease of communication, uniformity, group unity, group identification) but by placing (primary) standardisation efforts within a framework of modernity.

In modernity one strives towards order and the main aim of modernity is to create order out of disorder/chaos.(Blommaert et al. 2012: 15). Order equals normality, which means that an ordered person is someone who behaves according to the norms of society. (Ibid: 6). Norms of society are formulated in laws; regulations etc. and are upheld and dictated by the state. Consequently, in order for an idiom to reach the status of language it has to be ordered. This is done by codifying the language through dictionaries and grammars, thus promoting a colloquial means of communication (a mere language variety) to the status of a (national) standard language. Often, this kind of language standardisation entails language purification <sup>94</sup>. When a given language variety is subject to such standardisation, the language planners seek to purify it, so that it may become, in its normative standardised form, the 'cleanest', most authentic expression of the community (nation) that uses it. The desire for purity can also be explained within the framework of modernity, because as Blommaert et al note, closely related to the notions of order and normality is the notion of *purity* (Blommaert et al. 2012: 6).

A language is considered *impure* as soon as it is an overt mixture or hybrid. The notion of purity is thus closely connected to the notion of authenticity and originality. By codifying the language, order is achieved and as a natural consequence a norm (i.e. normality) is formulated. In the general wish for order (i.e. normality and purity) one must assume that the modernistic codifiers strive to include what they see as the originality of the language and discard the obvious hybrids, mixtures, interferences etc. from other languages and language varieties. (Blommaert et al. 2012: 5)

---

<sup>94</sup>Language purification is more commonly known as (linguistic) purism, cf. "2.2.3 Purism"

The (ordered, pure and normal) language, in turn, serves as an identity marker; if a person speaks a normal language (i.e. pure, uninfluenced by other languages) he also has a normal and orderly identity. He is not a mixture or hybrid, but pure and authentic. So, as a consequence of the creation of a group identity, markers of commonality with other groups are erased. In other words Ausbau (build-up) has as its aim to create Abstand (distance).

The need to obtain or maintain an unambiguous identity motivates measures against linguistic change which reflects hybridization, mixing with the other. The need to feel as a pure, orderly, normal member of a society may also explain why sometimes planned linguistic innovations are so quickly accepted into the broader media and by language users, which was the case in Croatia in the late 90's and at beginning of this century, according to Kordić when a 'purification' aimed at purging Croatian of so-called serbisms thrived. (Kordić 2010: 40)

The fear of losing one's identity when facing globalisation (the erasing of distinctiveness and assimilation into a larger group) may also explain the apparent opposition to Anglicisms (Blommaert et al. 2012: 15).

Bearing in mind that the aim is to strengthen group identity, it comes as no surprise that motivations for language planning may very well be a wish to be different from another group (nationality). As Laycock (2001), an expert on Melanesian linguistic diversity, suggests it, linguistic diversity may "in large measure [be] a partly conscious reaction" which reflects an attitude and wish for division as opposed to unity. (Laycock 2001: 171)

So, the diversification of otherwise similar language varieties may very well be, albeit initiated by historical and geographical circumstances, continuously intensified by speakers of different speech communities, who wish to be as different as possible from their similar counterpart. This is reminiscent of the view of Peti-Stantić in her definition of the secondary standardisation which has a separating rather than uniting function. (cf. p. 61)

So, once more we may conclude that a language is planned not only to ease communication but also in order to promote unity among its designated speakers and demote unity with speakers from whom there is a desire to be different.

Kaplan, Baldauf (1997) give an outline of a contextual basis for language planning. After having emphasised the importance of relating language planning to other kinds of Human Resource Development Planning



within the nation-state, they list four categories of actors who do the actual planning. 1) Governmental Agencies, i.e. Ministries, 2) Education Agencies (national, state and local), 3) Non/Quasi Government Organisations (Civil Service, Courts, Language Agencies, Religious organisations, Hospitals, etc. and 4) Other Organisations (for example Post Office, national TV-stations, big companies<sup>95</sup>). Following Kloss (1969) Haarmann (1989) and Hornberger (2006) I will add a fifth category, 5) individuals.

The scope of language planning is vast and involves not only the planning of the national and or official language, but also very much planning of minority languages. In this dissertation the focus is on the national (official) standard languages of Serbia and Croatia, resp., so I will not pay particular attention to matters concerning the large number of recognised minority and regional languages in the two countries (Croatia 22, Serbia 15)(Dobek 2013: 22; European Council 2013: 42).

Whether a linguistic community, (a society) should pursue similarity or dissimilarity with its neighbours is a question which in my opinion is instrumental in language policy-making and therefore in language planning as well, because what is regarded as a language is always ultimately a *political* question.<sup>96</sup>

### 2.2.2 One language, one nation

‘A nation’s life becomes evident through its language.’

‘The language is the only natural voice of the spirit of the nation.’

These were some of the dominant thoughts on the role of language (paraphrased by me), which were best formulated by the influential preromanticist German philosopher Herder (1744-1803), (Sundhaussen 1973:

---

<sup>95</sup> Coca-Cola is a good example. Coca-Cola Hellenic Bottle Company published in 2011, as the first company ever in Croatia, a manual on how to use Croatian in business communication (Hudeček et al. 2011). The aim of the manual is ‘to raise the level of linguistic culture in business communication’ and it has three chapters entitled “The Administrative Functional Style” p.10-26, “The Influence of English on Croatian” p. 27-57, “Other frequently disregarded language rules” p. 58-131 and a dictionary, p. 132-251, which lists all the wrong ways of spelling words along with undesirable words. This list includes a large number of Anglicisms. The manual was very much appreciated by the normativists in Croatia and received praise and recognition by the Croatist periodical *Jezik* in 2011 and 2012 and was awarded a prize by the EU in 2013.(Coca-Cola HBC Hrvatska 2016; Ham 2011b: 193–94)

<sup>96</sup> This does not mean that it, objectively, *is* a language but merely that it is regarded as such.

125). For instance: "Only through a language does a nation come into existence" or "Those who are brought up in the same language, belong to the nation of this language." (Sundhaussen 1973: 26, 125; 2009: 93). The idea that you need your own language in order to claim an own nationality was widespread in the 19<sup>th</sup> century nation building projects. As Joseph (2006: 22) points out, it is even implied in the Bible.<sup>97</sup>

In 1835, in the periodical of the Illyrian movement *Danica*, Ljudevit Gaj put it this way: "*Nima domorodztva prez lyubavi materinzkog jezika*" ('No patriotism without love for the mother tongue') (Sundhaussen 1973: 125). In short: If you wanted to claim nationhood, you needed a language to back it up.

Sue Wright (2004) provides a helpful categorisation of European countries defining them as either *state nations* or *nation states*. A state nation is a state as e.g. France where the borders of the state were defined before ideas of national identity had gained influence. Subsequently, the people(s) within the state were encouraged to embrace one common national identity, in this case the French. A nation state, on the other hand, is a state that has been built up *after* the nation has already been defined (Wright 2004: 26). The statehoods of Serbia and Croatia in 19<sup>th</sup> century were both conceptualized on the idea of a nation. So, a sense of national belonging came before state, not the other way around. Thus, both Serbia and Croatia may be defined as *nation states*. In Serbia in the mid-19<sup>th</sup> century, the modern ideas of building a state and a nation, were forced upon the people by the government (Sundhaussen 2009: 89). In Croatia in the mid-19<sup>th</sup> century, the politicians in the Croatian parliament *Sabor* tried, unsuccessfully, to gather the Croatian lands of Dalmatia, Slavonia and Croatia proper in one unit, first under Austrian rule (1849), then under Hungarian rule (1861). (Mønnesland 1999: 84–85)

Surely, there was a notion of nationality or national belonging in the peoples before language was deemed the crucial factor but, by becoming the crucial factor in defining nationalities, what was and what was not a separate language became increasingly important and defining a language became an important part of defining a nation. As we saw above ( p. 20), the thought that one language equals one nation and therefore one state, was so strong that Serbo-Croato-Slovenian was proclaimed the official language of the first Yugoslavia.

---

<sup>97</sup> Gen. 10:5: From these the coastland peoples spread in their lands, each with his own language, by their clans, in their nations (English Standard Version Bible)

The idea that you may lose your national distinctiveness, if you lose your language, or that your national distinctiveness will deteriorate if your (national) language deteriorates is very much alive today and may be seen as the main reason for the diligent efforts in the 1990's Croatia to remove what was considered Serbian lexemes as well as a few morphological and syntactic features, as outline in section 1.2 above. It is not only perceived Serbian influence that is in the spot-light in Croatia. There are also many voices which object to the innovations in the language usage induced by the contact with English as the language of the global world. Indeed, some internationalisms have probably been rooted out not because they are internationalisms, but because they are internationalisms used by Serbs. The objections to innovations induced by contact with so-called global languages are present among both Croats and Serbs, including the normative language standardisation bodies of Croatia and Serbia.

Above, in 2.2.1, we touched upon the notion of purism which is the term used to denote the efforts to purify and protect a language from foreign influence. In the next subsection we will take a closer look on this phenomenon and exemplify its various manifestations by referring to Croatian and Serbian puristic tendencies.

### 2.2.3 Purism

Purism has existed at least as long as the idea of national languages has. (Granić 2009: 65). Purism is a type of language correction which consists of regulating language users so that they correct their language and use a 'purer' vocabulary, morphemes or constructions when expressing themselves in said language. This is, for instance, the case when normativists in Serbia and Croatia alike warn speakers not to use nouns such as *menadžment* ('management') or *afinitet* ('affinity') and recommend the use of the domestic equivalents *uprava* and *sklonost* instead. Agents of purism tend to refer to the corrections as purification, authentication, indigenisation and traditionalisation of language (Neustupný 1989: 212) but we should not lose sight of the fact that purism presupposes that foreign influence is identified. Purists, professional linguists and laymen alike, often target the easiest identifiable contact-induced changes in a language – loanwords. I believe, as indicated in the introductory chapter that the subtle changes as for instance, innovative ways of forming compounds of two nouns or in combination with a semiword or the increased usage of the numeral *jedan* ('one') used as an indefinite article are more likely to 'go under the radar' and enter the usage norm without being

hindered by puristic tendencies – a claim and a belief that I will return to later.

Joseph (2006), who focuses on the politics of language, considers the motivation for language purification or purism to be rooted in a desire to protect a culture (i.e. nationality) from danger: “Multilingualism, language change and non-standard usage all feel like threats to the very foundation of a culture, since the language itself is the principal text in which the culture’s mental past and its present coherence are grounded.”(Ibid: 33)

George Thomas, in his book *Linguistic Purism* (1991), also emphasises the emotional reasons for purism, though in a more subdued way than Joseph, as he states that “[p]urism is the manifestation of a *desire* on the part of the speech community (or some section of it) to preserve a language from, or rid it of, putative foreign elements or other elements held to be undesirable” (G. Thomas 1991: 12) Thomas even indicates, by applying the adjective “putative” that he doubts whether some of the targeted foreign elements in fact are foreign elements.

The undesirable elements are labelled as incorrect in various ways. In the early puristic work of, for instance, Rožić (1904) the undesirable elements are called barbarisms and undesirable words and derivational affixes are called ‘monstrous’ (*nakazna*), ‘bad’ (*rđava*) ‘not good’ (*nije dobro*) ‘unnecessary’ (*nepotrebna*) or the mere mention of the model language (German, Turkish etc.) followed by the recommended word indicates its inappropriateness. Annamalai (1979; 1989) who defines purism as “[the] closure of non-native sources of [linguistic] enrichment” (1979: 36) shows that linguistic purism is focused on form rather than meaning (1989: 229). In other words, purism is an activity that serves to enhance the symbolic function of the purified language rather than the communicative one, albeit purists may claim that communicative efficiency is exactly why they support linguistic purism.

Later, typically when the language planning has been institutionalized and/or professionalized (G. Thomas 1992: 178), the corrective discourse becomes more moderate as for instance in the on-line language advisory service at the Institute for Croatian Language and Linguistics (IHJJ) where it is recommended not to write ‘web pages’ as a [N N]<sub>N</sub>-compound *internet-stranice* but rather by adjectivising the first noun (*internet>internetske*) thus forming the NP *internetske stranice*, where after it is pointed out that it is even better to use the domestic

lexeme *mreža* ('net') instead of *internet* thus forming the recommended NP *mrežne stranice* (IHJ 2015a) <sup>98</sup>.

Whether you view purism as an evil, a necessary evil or a (pure) necessity, there seems to be consensus among most scholars as to the goals of purism and techniques applied to purify a language. I hold that purism is extremely pertinent when dealing with language planning and policy and contact-induced language change because it "together with standardisation and enrichment [...] is key element of any language renewal movement" and "It provides a direct response to problems arising from language contact (especially where closely related languages are involved)." (G. Thomas 1992: 177)

Agents of linguistic purism aim at keeping the national language free (clean) of foreign and dialectal influences. Particular linguistic elements (lexemes, morphemes or syntactic phrases) are prescribed whereas others are proscribed. (Jakobsen 2010: 110–11) Purism may be directed at all linguistic levels but primarily targets the lexicon (G. Thomas 1991: 12). The main way in which this maintenance of the linguistic expression is done, is by substituting the unwanted element with a 'cleaner' one. This cleaner and thus more correct form of the element is generally proposed by linguists.

As mentioned above, puristic efforts are mostly targeted at immediately recognisable foreign elements such as foreign words (i.e. replicated lexical matter) but they may also be targeted at calques or translation loans, (i.e. replicated lexical pattern), especially when these calques introduce patterns in word-formation which are either foreign or in very restricted use in the replica language (i.e. motivates a change from a minor to a major use pattern). Purism may be described according to what phenomena it targets but it is also categorised according to the motivation and or the techniques applied in the puristic efforts.

Per Jakobsen (2010) lists four types and Thomas (1991) has an additional three. 1) *Archaising*, 2) *Ethnographic*, 3) *Reformist* and 4) *Xenophobic* (Jakobsen 2010: 112; G. Thomas 1991: 76–81) and 5) *Elitist*, 6)

---

<sup>98</sup> Original: "Takvu svezu dviju imenica bolje je zamijeniti svezom *pridjev + imenica* (*internetske stranice*, *internetski kafić*), koja je prihvatljivija u hrvatskome standardnojezičnom sustavu. Napominjemo ipak da je naziv *internetske stranice* bolje zamijeniti nazivom *mrežne stranice*."

*Playful* and 7) *Anti-purism* (G. Thomas 1991: 76–81)<sup>99</sup>. All seven of them are in play in the language planning and policy in either Croatian or Serbian.

*Archaising purism* is when the language planners seek to revive a perceived Golden Age of the language by (re)introducing lexemes and root morphemes, giving them new semantic content as a defence against foreign intrusion, e.g. the noun *računar* derived from the verb *računati* ('to calculate, to add up') has the modern meaning: 'computer', but originally had the meaning: 'book keeper'.

*Ethnographic purism* is based on the idealization of the rural language of the common folk. The language spoken in the countryside is perceived as being purer than that of the cities. The language reform of Vuk Stefanović Karadžić employed this kind of purism. The ethnographic purism entails that what is not considered part of the pure language of the countryside should be avoided.

*Reformist purism* is when language planners seek to break with a (political) past through closing the language to influences and loans from a language associated with this past. It could be exemplified the attempts of Hungarian or Croatian language planners to rid the language of German words in the 19<sup>th</sup> century or indeed the efforts by Croatian language planners to remove all that could be perceived as Serbian from the Croatian vocabulary, or at least from the Croatian standard language.

*Xenophobic purism* might be the most straightforward purism there is as its planners seek to eradicate, avoid or replace all foreign elements. It is undoubtedly xenophobic purism that make scholars such as Jakobsen (2010), Kordić (2010) and Clyne (1997) claim that purism may be linked to totalitarian regimes (Clyne) nationalism (Jakobsen and Kordić), racism and general xenophobia (Kordić).

*Elitist purism* may be regarded as the opposite of ethnographic purism because it seeks to purify of anything that is not part of the standard language prescribed on the basis of good writers. (Clyne 1997: 490) Croatian attitudes to language planning in the 19<sup>th</sup> century as well as the attitudes of Vuk's opponents in Serbia could be described as elitist purist. Whereas Vuk used the German Linguist Adelung's "Schreib, wie du sprichst!" ('Write as you speak!') as a motto, thereby including as many language users as possible, the Croat 20<sup>th</sup> century linguist Ljudevit Jonke,

---

<sup>99</sup> In a later work on lexical purism in Yugoslavia, Thomas (1992) only lists one type more than Jakobsen, the elitist type.

opposing the undesired consequences of Vuk's motto, promoted the motto "Piši kao što dobri pisci pišu!" ('Write as good writers do!'). Elitist purism maintains or paves the way for diglossia in a speech community; by forming the standard language according to the "good writers" and thus giving it the prestige of the elite, the difference between the colloquial and the standard language is enhanced, making the standard language more difficult to master. The difficulties in mastering the prestigious language variety – the standard language, in turn, make it the language of the elite.<sup>100</sup>

*Playful purism* refers to individuals' playful inventions in their vocabulary, mostly replacing foreign words, but these neologisms rarely if ever enter the norm (G. Thomas 1991: 80). In Serbia, as a reaction to the revival and renewal of distinctly Croatian vocabulary, a particular branch of playful purism arose. Serbs made Croatian mock-neologisms as for example *okolo-trbušni hlačo-držač* (an around-the-belly-trouser-holder) as a replacement for *kaiš* (a belt), a Turkish loanword.

*Anti-purism* is according to G. Thomas (1991) referring to Brozović (1970) a puristic reaction to purism. Brozović argues that anti-purism is very similar to purism because its supporters are very intolerant and 'there is a danger that a new antipurism jeopardizes the positive, stabilised, absorbed, useful and well-functioning results of the old purism'<sup>101</sup> (Brozović 1970: 79). Brozović does not go into details about who has done what to which positive puristic results of the past, but it is likely that he is referring to the marginalization of a distinctly Croat vocabulary in favour of the so-called federal vocabulary as described by Pranjković (2006: 45–46).<sup>102</sup> A more contemporary example of anti-purism, is what we see when Serbs object to neologisms (calques) made in Serbian, in Serbia by Serbs because they view this particular strategy as Croatian, which is why for instance *nadstolnik*<sup>103</sup> a proposed substitution for 'desktop computer' was never accepted (Granić 2009: 77).

---

<sup>100</sup> Cf. Joseph (2006: 19) on how elitist language serves to separate the sheep from the goats.

<sup>101</sup> Original: ...javlja se opasnost da jedan novi antipurizam ugrozi i neke pozitivne, tj. već ustaljene, apsorbirane, kvalitetno korisne i dobro funkcionalne rezultate toga staroga purizma.

<sup>102</sup> Also mentioned in subsection 1.1.5.

<sup>103</sup> *nad* ('above') *stol* ('table') *-nik* (nominalising formant) > *nad-stol-nik*

Omitting playful purism and anti-purism because they are not exclusively externally or internally oriented, G. Thomas categorises xenophobic purism as the unmarked, and thus most commonly identified form of purism, and as the only type of purism which has an external target. The remaining types: archaising, ethnographic, reformist and elitist all have an internal perspective. (1991: 81)

I would rather say that whereas xenophobic purism is an attack on all things perceived as foreign in a language, the other types represent the sources in which the replacements of the foreign elements are found. In other words, practised purism is never only one of the above, but rather a combination of xenophobic purism and some of the others (still excluding playful purism and anti-purism).

In my following paraphrase of the words of the Croatian language guide from 1999 (*Hrvatski jezični savjetnik*), the puristic attitude of today's leading Croatian linguistic scholars is clear: 'The majority of Croatian linguists agree to respect the centuries-old tradition of Croatian linguistic purism and in principle support the puristic demand not to use unnecessary loanwords, but substitute them with good domestic words.' (Barić et al. 1999: 106)

Croatian purism has in practice been a combination of all five purisms; xenophobic, elitist, ethnographic, archaising and even reformist purism<sup>104</sup>. Xenophobic, because there is a tendency to deem all perceived foreign matter (lexemes and morphemes) and patterns (syntactic, word-formational, phrasal) less correct or even "not in the spirit of the Croatian language" and in lieu of foreign linguistic matter and patterns promote existing, new or archaic inherited linguistic matter or patterns. In other or the same instances the recommended linguistic matter is taken from what is considered older canonical Croatian literature or usage norm. The reinstated matter, however, also frequently exists in dialects, thus making it an instance of ethnographic purism. And, as we have seen, the xenophobic purism may be directed at one or more particular languages belonging to speech communities with which the Croatian speech community has, in the period in question, an undesirable past, e.g. German in the 19<sup>th</sup> century, Serbian, in the late 20<sup>th</sup> and early 21<sup>st</sup> centuries.

---

<sup>104</sup> Thomas would not agree in this as he sees archaising and reformist as well as elitist and ethnographic purisms as each other's contradiction and thus allows only for three-dimensional purisms (G. Thomas 1991: 82–83)



Serbian purism is less vigorous than its Croatian counterpart. The common attitude among contemporary Serbian linguists towards loanwords, the most salient foreign influence in a language, is often characterized as non-purist (Plankoš (ed.) 1996, 24; 46; 55; 317). Klajn (2008: 157) even sees the *antipuristic* tendencies in Serbian of today as a counteraction to Croatian purism thus identifying the tendency to replicate foreign linguistic matter as an anti-Croat tendency among Serbs, that is, as a wish to differentiate themselves from Croats. So in this respect, replicated linguistic matter is definitely a factor which divides Croats from Serbs or rather Croatian from Serbian.

Vuk Stefanović Karadžić, on the other hand, pursued a reformist and ethnographic purism when he in the first half of the 19<sup>th</sup> century omitted the Russian and Church Slavonic linguistic matter (lexis) and patterns (constructions) present in Slaveno-Serbian from his reformed colloquial-based Serbian. The motivation for omitting these features is primarily founded on ethnographic puristic tendencies and cannot be categorised as xenophobic because other foreign linguistic matter (lexemes and morphemes from Turkish), which was part of the colloquial rural language usage was codified as part of the Serbian language corpus.

Xenophobic purism in Serbian existed in the 19<sup>th</sup> century but remained, according to Thomas (1992: 180) referring to Herrity (1978) and according to Fekete (1996a: 156) largely inconsequential. If not xenophobic then reformist purism exists in contemporary Serbian Language Planning. Calls for stopping the onslaught of foreign words<sup>105</sup>, linguistic conferences dealing with issues concerning foreign influence on the development of the Serbian standard language<sup>106</sup> and (re)actions from *The Board for Standardisation of the Serbian Language* show an enhanced awareness and wariness of contact-induced language change of Serbian.<sup>107</sup>

Normativists in today's Serbia sometimes point out the undesirable in foreign influence, i.e. contact-induced linguistic innovations, e.g. Klajn (Piper and Klajn 2013: 355) and Brborić et al. (2006: 101; 391) and therefore could be labelled *moderate* reformist purists as they accept the

---

<sup>105</sup> Medić, M. (2001) *Naježda stranih reči na srpski jezik*, Nolit, Beograd

<sup>106</sup> *Konkurencija jezičkih sredstava u srpskom jeziku* (1999), *Intralingvistički i ekstralingvistički činioci u formiranju i razvoju srpskog standardnog jezika* (2012), *Inovacioni procesi u srpskom književnom jeziku - uticaji drugih jezika i kultura* (2013) organised by the University of Belgrade.

<sup>107</sup> The recommendations from *The Board for Standardisation of the Serbian Language* are revisited in subsection 4.2.1.

necessity and inevitability of foreign influence, but at the same time recommend caution and the use of domestic equivalents when possible. The conclusion in an advice on the use of foreign terminology given by *The Board for Standardisation of the Serbian Language* sums up this seemingly ambivalent attitude: 'Foreign terms should be translated – when and where it is possible and sensible to do so.' (Brborić et al 2006: 101)

*A few examples of puristic recommendations by language planners in Croatia and Serbia*

a) A primary recommendation of a domestic lexeme rather than a foreign one: \**afinitet* > *sklonost* ('affinity')

b) A primary recommendation of a domestic lexeme rather than a foreign one, secondarily an international foreign word of Latin or Greek origin (internationalism) is recommended over an English loanword:

\**donor* > *donator* > *darivatelj* ('donor')

c) A domestic prefix (morpheme) is recommended over an international: \**multinacionalni* > *višenacionalni* ('multinational')

d) Domestic NPs equivalent to English compounds as 'web pages', 'contact group', 'fitness centre' or 'taxi driver' is recommended over the undesired domestic compound formation, i.e. a hyphenated [N-N]<sub>N</sub>:

d1) \**internet-stranice* > *internetske stranice* > *mrežne stranice*

[N-N]<sub>N</sub>

[A N]<sub>NP</sub>

[A N]<sub>NP</sub>

d2) \**taksi-vozač* > *vozač taksija*

[N N]<sub>N</sub>

[N N]<sub>NP</sub>

or an unhyphenated [N N]<sub>N</sub>:

d3) \**fitness centar* > *centar za fitness*

[N N]<sub>N</sub>

[N PP[P N]]<sub>NP</sub>

d4) \**kontakt grupa* > *kontaktna grupa*

[N N]<sub>N</sub>

[A N]<sub>NP</sub>

e) A domestic NP (as in d1 and d4) or an apposition equivalent to the English: 'Stanford University':

\**Stanford Univerzitet* > *Stanfordski univerzitet* / *Univerzitet Stanford*

[N N]<sub>N</sub>

[A N]<sub>NP</sub>

N

N

f) Advice against the use of the phrase *Moje ime je* ('My name is') as it is a result of English influence instead of the domestic *Ime mi je, Zovem se* or *Ja sam*:

*Moje ime je*\* > *Ime mi je / Ja sam / Zovem se*

'My name is , Name me is / I am / (I) call myself'<sup>108</sup>

How the language planning institutions in Croatia and Serbia react to specific contact-induced language changes and how much effect their reactions have on the usage norm will be dealt with in further detail concurrently with the investigation of the concrete examples of linguistic innovation in chapter 4, p.146ff.

As do many scholars in the field, I regard purism as a reaction and an attempt to rid a standard language of its perceived foreign elements.<sup>109</sup> I will also take a closer look at what is defined as the most undesirable elements (lexemes, morphemes, word order, word-formation, syntactic features), whether it is important whence these perceived foreign elements come (English, Turkish, German, etc.), what their recommended substitutions consist of (existing, newly formed or hitherto archaic 'domestic' elements), with which intensity the correction of faulty language use is pursued, by whom it is pursued and lastly what triggers linguistic purism.

Purism is, however intriguing, only one part of language planning, so we will now return to the greater scope of language planning and through the notions of status and function proceed to look at four suggested dimensions of language planning: *Corpus planning*, *Status planning*, *Acquisition planning* and *Prestige planning* which will lead to a preview of how the planning of the standard languages Serbo-Croatian, Croatian and Serbian have been done. This preview will show how these specific standard languages came about, and also use these specific cases to illustrate some general regularities of language planning.

---

<sup>108</sup> Recommendations (a-d1) are retrieved from the Croatian on-line language guide from Institute of Croatian Language and Linguistics (IHJ) 2015a; 2015b; 2015c; 2015d). d2-3 is from the Coca-Cola language manual (Coca-Cola HBC Hrvatska 2016: 97) Recommendation (d.ii) is a from *The Board for Standardisation of the Serbian Language* in (Brborić et al 2006: 54), Recommendation (f) was retrieved from both 'The Normative Grammar of Serbian' (Piper and Klajn 2013: 355) and *Srpski ili anglosrpski – pitanje je sad* ('Serbian or Angloserbian- that is the question') (T. Prčić 2006: 416–17) who is also the source of recommendation (e) (Ibid.: 415)

<sup>109</sup> G.Thomas (1991: 72) also includes purification of what by some purists are considered inappropriate neologisms and dialectalisms.

### 2.3 Language status and language function

When addressing the issue of language planning and policy one cannot omit the *status* and *function* of a language, i.e. what it takes for an idiom to have the status of language and what functions an idiom needs in order to be considered a language.

Ammon (1989) in his endeavours to outline a descriptive framework for Status/Function of languages takes Heinz Kloss' well-known distinction of Ausbau and Abstand-languages as a point of departure. Ausbau-languages are languages by virtue of their cultivation (elaboration, modernization). Abstand-languages are languages solely by their distance to other linguistic systems. Ammon aptly renames the latter *dissimilar languages* to avoid the assumption of geographical distance. For instance: the geographical distance between Serbian and Albanian is not greater than that between Serbian and Macedonian is. Nevertheless Albanian and Serbian are Abstand (distance > dissimilarity) languages whereas the relationship between Serbian and Macedonian is one of Ausbau (cultivation).

As mentioned briefly in chapter 1, Heinz Kloss (1967) originally showed that an idiom, on the one hand can be defined as a language on account of its dissimilarity (Abstand) to other linguistic systems and on the other hand an idiom may also gain the status of a language through cultivation, i.e. elaboration, to meet the communicative and symbolic needs of its speakers (Ausbau). Through this cultivation the idiom gains in functionality and at the same time loses similarity to idioms spoken by speakers within the same language family (e.g. Macedonian and Serbian). Kloss exemplifies Ausbau-languages through the relation between Danish and Swedish, Czech and Slovak, Icelandic and Faroese, each pair of languages having been developed out of one preliterate linguistic community. The Serbian-Croatian situation is somewhat different than the Danish-Swedish, Czech-Slovak, and Icelandic-Faroese one because both idioms had at least one literate linguistic community before their (attempted) unification in the 20<sup>th</sup> century. Another defining, and related, difference between the internal relation in the mentioned language pairs and the Serbian-Croatian relation is, that standard Serbian and Croatian have been cultivated (ausgebaut) on the same foundation, that is, the base dialect Neoštokavian whereas e.g. standard Czech and standard Slovak are cultivations of different base dialects (Kloss 1967: 31). As the separate cultivation of the Neoštokavian dialect had occurred and still occurred around two geographical centres at the time, Kloss used the

term ‘a polycentric standard language’<sup>110</sup> and applied it to Serbo-Croatian (having been developed around two centres at the time), English (in America, in Great Britain.), Portuguese (in Portugal, in Brazil) and finally Romanian and Moldavian (in Romania, in the USSR) (Kloss 1967: 31–32). The fact remains, though, that separate cultivation (Ausbau) of Croatian and Serbian be it before, during or after the official life time of standard Serbo-Croatian, leads to a loss of similarity, regardless of the shared base dialect.

The Ausbau-argument applied when advocating Croatian’s dissimilarity to Serbian is exactly the separate cultivation or different *acculturation* called either *civilizacijska*, *kulturna* or even *kulturno-civilizacijska nadgradnja* of the two languages. Brozović, a proponent of designating Serbian and Croatian as one standard but with standard versions (variants) argues, while comparing the internal differences to those between British and American English, that the differences lie in the different ways in which the basic standard (i.e. the standardised base dialect) has developed into different variants through time, or as he puts it: *the relation between the dialectal raw basis of a standard language and its acculturated versions*<sup>111</sup> (Brozović 1985: 82–83). It is, in other words, the different acculturations (elaboration, modernisation) that are responsible for the different manifestations of the basic standard and it is these different manifestations<sup>112</sup>, these *language varieties* that by Brozović (1985) are called *variants*, and which were later officially promoted into *languages*.

### 2.3.1 Status, Corpus, Prestige and Acquisition Planning

In this subsection we will go through four labels applied to different aspects of language planning. Language *corpus* planning and language *status* planning are terms coined by Kloss (1969: 81) whereas the notions of *prestige* planning and *acquisition* planning have been introduced into the language planning discussion at a later date by Haarmann (1989) and Cooper (1989), resp.

---

<sup>110</sup> A term coined by William A. Stewart and first seen in a publication from 1968 ([Stewart 1968]). Also known as „pluricentric languages“

<sup>111</sup> Original: „odnos[u] između dijalektne, sirovinske osnovice standardnoga jezika i njegove civilizacijske nadgradnje“

<sup>112</sup> Original: „različiti konkretni realizacijski oblici (Brozović 1985: 84)

### ***Corpus planning***

According to Kloss (1969: 81) *corpus planning* denotes when an active agent (an agency, group of individuals or one individual) endeavours to modify the existing language itself by setting up standards for or altering existing standards for which writing system is to be used, how the language is to be spelled, how its wordforms are to be inflected and how the word-order should be, and also carries out the task of developing or coining new specialized vocabulary .

The list of features which may be subject to modification can be prolonged as for instance Radovanović does it, adding orthoepy (pronunciation), phonology, syntax, semantics, word-formation (derivation), lexical innovation (including adaptation of foreign lexemes), structuring of text, discourse and speech acts, style (genre) and pragmatics (interaction). (Radovanović 2004: 133–34). Lexical innovation should be further subdivided into the categories of modernisation and renovation, referring to innovations that fill a so-called lexical gap and renovations that represent new features (expressions, lexemes, morphemes, etc.) that take the place of existing ones (Hornberger 2006: 32)<sup>113</sup>. These renovations may, but need not, be the result of linguistic purism.

All the mentioned features existing in a language, i.e. in the corpus of a language, may be the object of planned modification from specific agents and the means of such corpus planning are usually in the form of manuals (dictionaries, grammars, guides) which the speaker, who wishes to use the language according to the standardised norms, can consult or which are used as learning material in language training in, e.g. schools.

### ***Status planning***

*Status planning* concerns itself with the status of a particular language within a society, mostly in legal terms. A language can have the status of official (national) or minority language.

Ammon (1989) develops the notion of language status and language function and suggests that the status a language equals its social position in a country, meaning that if a language is the official language of a state it has a very high status. This does not necessarily mean that it

---

<sup>113</sup> Within the field of lexical borrowing, Jørn Lund, 2003 cited in Gottlieb 2014 calls these renovations *cuckoos*, because they enter the language and exist alongside the existing term, but after a while they "push the former rulers out the nest." (Gottlieb's translation from Danish to English).

has a high degree of function, because a language variety may enjoy official status without being used by very many and/or by the highest ranking members of the society (Ammon 1989: 26). In other words; official status and functional status are not the same.

Radovanović (2003; 2004; 2009) sees the evolution of a standard language as ten steps of language planning, where the four steps 5) *the officialisation*<sup>114</sup>, i.e. official ‘acceptance’ of a language variety by the authorities, 6) its *implementation*, i.e. its actual usage 7) its *expansion* into new geographical and social areas and 8) its propagation via education (of both children and adults), publishing houses and mass media, which he calls ‘*cultivation*’, concern status planning (Radovanović 2009: 205). Of the remaining six steps five are considered as belonging to corpus planning and one: *selection* is considered to be the link between the two types of planning (Ibid.). We will return to Radovanović’s ten steps of language in section 2.4, below.

### ***Prestige planning***

Haarmann (1989; 1990) suggests a third dimension of language planning: *prestige planning* which is introduced to account for the fact that in the language planning processes it is not only important to gain *acceptance* of the planned language variety or planned changes of said language variety. It is also important “*Who accepts what planning provisions from whom and under what conditions*” (Haarmann 1990: 123) or in Kaplan and Baldauf’s words: how status and corpus planning activities are acted upon by *actors* and received by *people* (Kaplan and Baldauf 1997: 50). Haarmann focuses on what Radovanović calls *acceptance* and *implementation* and emphasises that not only the planners and their products are important but that the degree of implementation done by language users is dependent on *who* the planners are and on *what* the social and cultural circumstances are for the recipient. The notion of prestige planning or rather the role of the planners’ prestige and influence in the successful implementation of the planned modifications and recommendations is, I believe, key to understanding why some planned measures are successfully implemented and others are not.

---

<sup>114</sup> an English term borrowed from Hornberger (2006), below in this section

### ***Acquisition planning***

Hornberger (2006) provides an integrative framework of language policy and language planning goals where she integrates the foci and terminology of various scholars<sup>115</sup> within the field(s) of Language planning and Language policy. Hornberger, like Haarmann, highlights an additional aspect of language planning - *acquisition planning* - and explicates that this dimension of planning is aimed at the *language users*, whereas status planning is aimed at *language use* whilst corpus planning is aimed at *language* itself.

*Acquisition planning* (originally introduced by Cooper (1989)) is similar to Haarmann's *prestige planning* as regards its focus on the language users, but where Haarmann includes both the planners and the recipients of the planning, acquisition planning is primarily concerned with the recipients and the conditions provided for them to receive (and subsequently accept and implement) the planning efforts. (Hornberger 2006: 32; Wright 2004: 61). Acquisition planning refers to activities pertaining a) acquisition of a language, be it a marginalised mother tongue or a foreign language and b) maintenance of an existing language. These planning activities primarily take place in the educational system, but may also be conducted through the media. E.g.: Croatia's state-run TV and Radio Broadcasting Company broadcasts at least two programmes where linguists offer advice on how to speak and write proper standard Croatian. The programmes that were on in January 2016 were *Govorimo hrvatski* ('We speak Croatian') and *Hrvatski naš svagdašnji* ('Our own everyday Croatian').

### ***Actors***

So, planning the status of a language is when an actor (an agency, group of individuals or one individual) endeavours to modify the legal and social position of a language (an idiom) within a society.

Whereas corpus planning is carried out by linguistic specialists, status planning is dealt with by legislators. Actions within corpus and status planning are, in spite of their different actors, very often intertwined. For instance in the case of Croatian, where several corpus planning works during the 20<sup>th</sup> century had, if not as their primary, then as a secondary goal the consolidation of the status of Croatian as a standard language, independent of Serbian and with a social positioning equal to

---

<sup>115</sup> Cooper (1989), Ferguson (1968), Haugen (1983), Hornberger (1994), Kloss (1968), Nahir (1984), Neustupny (1974), Rabin (1971) and Stewart (1968)



that of Serbian (thus denying the existence of Serbo-Croatian as a single standard language).<sup>116</sup>

Whether or not or how much effort is put into the implementation of a strategy of language planning is a question of political will. The *political will* is in some countries formulated in more or less detail in an official (state) language policy. In other countries the policies and applied strategies are a result of consensus within norm giving institutions, which are not necessarily a reflection of the attitudes of the state government. In the case of Croatian and Serbian, and before that Serbo-Croatian, there has been virtually no legally formulated state language policy regarding Serbian and Croatian other than the official name and alphabet. However, the norm-giving institutions as for instance the Serbian Academy of Science and Arts, the Croatian Academy of Science and Arts, the Matica Srpska, the Matica Hrvatska, the Board for Standardisation of the Serbian Language, the Council for Standard Croatian Language Norm (dissolved in 2012) and the Institute for the Croatian Language and Linguistics<sup>117</sup>, have played and play a major role in formulating rules and recommendations as well as in publishing normative works on what Standard Serbian or Standard Croatian is.

## 2.4 Planning of Serbian, Croatian and Serbo-Croatian

We will now by applying the terminology introduced above take a closer look at how language planning manifested and manifest itself in Serbian, Croatian and Serbo-Croatian in order to see who is planning and how and what they are planning. On the basis of Radovanović' ten phases of language planning which will be commented on and discussed through other general outlines and frameworks of language planning and exemplified with language planning efforts within Serbian, Croatian and Serbo-Croatian I will continue the discussion of what language planning is and supply concrete examples of it.

As mentioned earlier, behind and inextricably connected to language planning is language policy, or rather language policy-making. Kaplan and Baldauf (1997: 6) point to four different actors 1) Govern-

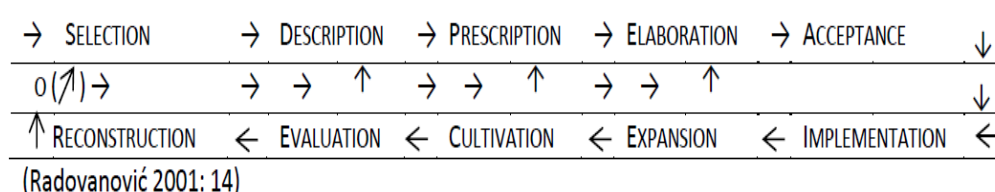
---

<sup>116</sup> Cf. for instance the overviews provided by Marko Samardžija and Ivo Pranjković in Samardžija 2006 p. 9-28, 29-57 resp.

<sup>117</sup> Original names: *Srpska akademija nauka i umetnosti*, *Hrvatska akademija znanosti i umjetnosti*, *Matica Srpska*, *Matica Hrvatska*, *Odbor za standardizaciju srpskog jezika*, *Vijeće za normu hrvatskoga standardnog jezika*, *Institut za hrvatski jezik i jezikoslovlje*.

ment agencies, 2) Education agencies, 3) non-/quasigovernment agencies and 4) other organisations. In addition to these four a fifth category '5) individuals (following Haarmann, Figure 2.4) may be added. These actors are responsible for initiating the planning phases as outlined by Radovanović in Figure 2.3.

*Figure 2.3 - Radovanović's ten phases of language planning*



Radovanović' 10 phases or steps are to be viewed as a circular process, where some phases can coincide (e.g. grammar books both describe and prescribe), and others may come in another order than put up here. (E.g. norms may be implemented before they are fully accepted: IMPLEMENTATION → ACCEPTANCE). The horizontal arrows in the first and third line of the model show the circular process whereas the 0 shows the typical point of departure in language planning. The vertical and horizontal arrows in the second line show how the different phases may be concurrent and the tilted arrow in the brackets (left) shows how the tenth and 'final' phase may lead to the selection of another base dialect which then leads to the 9 phases being recommenced.

*Figure 2.4 – Haarmann's ideal typology of language planning*

	Ranges of language planning		Ranges of language cultivation		
	Governmental activities	Activities of agencies	Group activities	Activities of individuals	
Language status					Language status
Language prestige	officially promoted	institutionally promoted	promoted by a pressure group	individually promoted	Language prestige
Language corpus					Language corpus
	Level 4	Level 3	Level 2	Level 1	
	Maximum	Efficiency in terms of organizational impact		Minimum	

(Haarmann 1990: 120)

Haarmann puts emphasis on whether the language is cultivated by mere individuals or groups, and whether these activities are supported and if

so whether they are supported by strong or less strong groups within the society. Only when the planning activities are promoted by institutions or the government of a given country, does he term it language planning and chooses to term it language cultivation when the planning is promoted by non-institutionalised groups and individuals. (right side) He also, not surprisingly, claims that the more substantial support a cultivation or planning measure gets, the more efficient the impact will become (horizontal two-pointed arrow at the bottom of the model). The arrows pointing vertically in two directions indicate that the planning of both status and corpus emanates from the various levels of prestige enjoyed by the planners/cultivators. Thus the significance of the prestige enjoyed by initiators and promoters of both language status and corpus activities is not only vital to but also proportionally connected to the successful implementation and acceptance of the language planning measures, whether these measures pertain to the status or corpus of the language.

Figure 2.5 – Hornberger’s integrative framework of language planning typologies

Types	Policy planning approach (on form)	Cultivation planning approach (on function)
Status planning (about uses of language)	<b>Officialization</b> <b>Nationalization</b> <b>Standardization of status</b> <b>Proscription</b>	<b>Revival</b> <b>Maintenance</b> <b>Spread</b>
Acquisition planning (about users of language)	<b>Group</b> <b>Education/School</b> <b>Literary</b> <b>Religious</b> <b>Mass media</b> <b>Work</b>	<b>Reacquisition</b> <b>Maintenance</b>
	<b>Selection</b> Language’s formal role in society <i>Extra-linguistic aims</i>	<b>Implementation</b> Language’s functional role in society <i>Extra-linguistic aims</i>
Corpus planning (about language)	<b>Standardization of corpus</b>	<b>Modernization</b> (new functions) <b>-Lexical</b> <b>-Stylistic</b>  <b>Renovation</b> (new forms, old functions) <b>-Purification</b> <b>-Reform</b> <b>-Stylistic simplification</b> <b>-Terminology unification</b>
	<b>Codification</b> Language’s forms <i>Linguistic aims</i>	<b>Elaboration</b> Language’s functions <i>Semi-linguistic aims</i>

(Hornberger 2006: 29)

As Hornberger’s model is concerned with planning not only of traditional standard languages but also includes auxiliary codes such as sign language, second languages and languages without a previous writing system I have deleted some of the headings in her original model (published 2006). Under the punctuated lines in Hornberger’s model we find Haugen’s terminology which approximately covers what is immediately above the punctuated line. Hornberger’s model serves to highlight the importance of not only the actors who promote the planning activities as

we saw it in Haarmann's model but the *recipients*: groups, education/school, literature, religion, mass media and work. These six labels cover the [language] users who "are targeted to receive opportunity and/or incentive to learn the given language" (Hornberger 2006: 32).

An example which illustrates the importance of prestige enjoyed by the actors and recipients in the acquisition planning is the manual and dictionary of Babić et al. which was banned in communist era, to be reprinted and published in Croatia in 1990, (cf. p. 26) emanated into the recipient labelled "work" when the Coca-Cola company in Croatia in 2011 wrote and published a manual with the subtitle 'business communication' in which the orthographic rules recommended and followed throughout this 272 pages long manual<sup>118</sup> are based on the Babić-Finka-Moguš orthographic dictionary (Ham 2011b: 194).

In the following I will use Radovanović' 10 phases illustrated in Figure 2.3 above as a point of departure for presenting and discussing language planning in Yugoslavia, Serbia and Croatia and when appropriate refer to some of the aspects illustrated in Haarmann's and Hornberger's models (Figures 2.4 and 2.5).

#### **2.4.1 Selection**

Selection is the key phase in the establishing of a standard language. In this phase an idiom (language variety, dialect, or sociolect) is chosen as the foundation in which the next phases in language planning are grounded. In the case of Serbo-Croatian (Serbian and Croatian) the Neoštokavian dialect of Eastern Herzegovina was selected (Radovanović 2003: 191). The selection of a language variety naturally predates the next phases as one can hardly describe nor prescribe the use of anything unless one has chosen what it is one wants to describe, prescribe etc.

The selection of the dialect was firstly done by individuals (Vuk Karadžić) and non-government organisations (The Illyrian movement), then the selection of Neoštokavian was promoted as a group activity, i.e. The Vienna Literary Agreement and by writers and poets in both Serbia and Croatia who chose to write in Štokavian, following the orthographic principles, described and prescribed in various normative works. Only later was their selected language variety along with further language planning steps accepted by the authorities. (Cf. subsection 2.4.4, below)

---

<sup>118</sup> cf. footnote 95, p.74

Selection of a base dialect was by the early language planning theoreticians (Haugen, Kloss) seen as a status planning phase, wherein a particular idiom is selected as the language variety which is to go through the ensuing corpus planning phases. Some, as e.g. Wright, even claim that this is the main purpose of status planning. “Status planning concerns itself with the choice of the varieties that will become the official language(s) of the state, in particular the medium of its institutions.” (Wright 2004: 43)

The significance of the selected base dialect has also been recognised by contemporary linguists who consider Croatian as a language in its own linguistic right as members of the Council for Standard Croatian Language Norm<sup>119</sup> (including its chairman Radoslav Katičić) claim that the base dialect of Croatian is *not* Neoštokavian but that Croatian was based not on one dialect but rather on a hybrid, of Kajkavian, Čakavian and Neoštokavian dialects which later, during the attempted unification of Croatian and Serbian was stylized in a Neoštokavian way. Were it to be broadly recognised and accepted that the base dialects of Serbian and Croatian are in fact not one and the same, Kloss’ distinction between a pluricentric language and two distinct Ausbau-languages as e.g. Czech and Slovak or Danish and Swedish, standard Serbian and standard Croatian would, solely on the basis of the their base dialects, be considered to be two separate languages. Other Croatian linguists (as e.g. Peti-Stantić 2008) do not dispute the fact that Serbian and Croatian share the same base dialect. However, she considers the selected base dialect to be of minor importance in language planning whereas the major role in language planning belongs to the *elaboration* of the language variety in question (Peti-Stantić 2008: 15)<sup>120</sup>

However, this view is not only disputed by scholars outside Croatia but also within including other members of this council (specifically Ivo Pranjković) because, as the argument goes, a standard language *must* have a base dialect, which in the instance of Croatian *is* Neoštokavian, and that it is in the stylization, i.e. the more superficial forming of the Croatian standard language, that Kajkavian and Čakavian features emerge, concretely on the lexical level (VNHSJ 2013: 114–19). So, despite individual promotion by scholars and institutional promotion by some

---

<sup>119</sup> Established by Croatia’s Ministry of Education and Sports in 2005

<sup>120</sup> Peti-Stantić, who uses the same line of argumentation as Brozović (cf. p 63 and p. 86), does not use the terms language planning and elaboration, but *standardisation* and *acculturation*.

normative bodies in Croatia to remove the perceived existence of a shared base dialect, the precondition for a shared standard language, most linguists consider the base dialect of all successor languages of Serbo-Croatian to be Neoštokavian, as it was, undisputedly, considered the base dialect of Serbo-Croatian.

The selection of the Neoštokavian East-Herzegovinian language variety in the 19<sup>th</sup> century was, as we learned in the Introduction, not done without hesitation and opposing arguments in both Croatia and Serbia. The selection was, however, done very deliberately and served the very clear political purpose of achieving a separate South Slavic statehood. Aiming towards a separate statehood was not unique for Serbs and Croats, but a rather common feature among the stateless European nations, also a point put forward by Wright (2004: 47), mentioned in subsection 2.2.2 above.

Radovanović himself (inspired by (Škiljan 1988)) proposes to ascribe the selection process a role which links the corpus and status planning, because it is the selection that is the initial phase which determines the course of the ensuing corpus and status planning and it is also in this phase that another language variety may be chosen (selected), thus setting off the planning of a new standard language (illustrated by the tilted arrow in Figure 2.3 above) (Radovanović 2004: 65; 2009: 205).

#### **2.4.2 Description and Prescription**

Ideally, description is just that; a full non-discriminatory description of the selected language variety through dictionaries, grammars and other linguistic studies. However, description is often done alongside prescription, which is a very discriminatory approach where the ‘proper’ way of using the language variety is prescribed as well as described. At the same time ‘improper’ or ‘irregular’ usage is proscribed. The scholarly and educational publications stemming from the description and prescription become the standardized, prescribed norm, which all users of this particular standard language are to follow (Radovanović 2003: 192–93).

The linguistic *description* of the selected language variety as well as its *prescription* was done initially by Vuk Stefanović Karadžić (1787–1864) and Đuro Daničić (1825–1882) and later by Vuk’s followers in both Croatia (Broz’ orthographic dictionary (1892), Broz and Iveković’ dictionary (1901) and Maretić’ grammar (1899)) and Serbia (Orthographic dictionaries by Čuturilo (1884) and Petrović (1914) as well as a grammar by Novaković (1894)).

Đuro Daničić is not as easily categorised according 'country' as the ensuing linguists because he was the first to supply his work with titles by which he affirmed the idea that Croatian and Serbian were one language. He was the initial writer of the monumental work "Rječnik hrvatskoga ili srpskoga jezika" (1880-1976) (Dictionary of the Croatian or Serbian language), which was initiated at The Yugoslav Academy of Arts and Sciences (est. 1866) in Zagreb. Albeit a Serb, he lived and worked in Zagreb 1866-1873 and again from 1877 to his death in 1882 and had a profound influence on the linguistic bonding between Serbs and Croats (P. Ivić 1998: 234–35).

In the endeavours to codify the standard Croatian and Serbian after, as well as before the Neoštokavian was selected, *purism* was a major influence. In the 19<sup>th</sup> century, the notion that a mixture or a hybrid of languages was impure and threatening to national revival was widespread. This was closely linked to the dominant romanticist philosophy of the time, which, among many other ideas, promoted the idea of purity originating in the Golden Age, and of cleansing language so that it might regain its former purer form. In the Serbian of Vuk Karadžić, Russian and Church Slavonic influences that had entered Serbian were rooted out and in Croatian a lot of German and Hungarian influences were deemed a danger to the pure Croatian expression. The purism of the time was therefore mostly reformist, among Vuk and his followers (both Serbian and Croatian) ethnographic/reformist, among Croats who were not Vukovites reformist/elitist whereas Vuk's adversaries (if at all puristic) mostly pursued elitist purism (cf. "2.2.3 Purism"). Purism is, as mentioned earlier, mostly directed at the lexis, but other linguistic levels, for instance the syntactic level, were also scrutinized. Participial use of gerunds were deemed unacceptable by Vuk Karadžić and a tendency to place the main verb last (a perceived influence from German) was treated likewise. The *proscription* of 'irregular' or 'inferior' words found its way into grammars (e.g. Maretić 1899) and dictionaries and there were even published dictionaries of the proscribed words and phrases alongside the proper words or phrases, e.g. Rožić (1904) *Barbarizmi u hrvatskom ili srpskom jeziku* ('Barbarisms in Croatian or Serbian').

In Croatia in the beginning of the 20<sup>th</sup> century and until 1921 the prescribed norm for spelling and orthography was set by Ivan Broz' *Hrvatski pravopis* (Zagreb 1892, edited by Dragutin Boranić from 1904) and the normative grammar used in secondary education until 1928 was Maretić's *Hrvatska ili srpska gramatika za srednje škole*. In lexicography,



the norm was set by the dictionary *Rječnik hrvatskoga jezika* by Iveković and Broz (first published in Zagreb, 1901) (Samardžija 2006: 9-10). Broz' *Hrvatski pravopis* was to become the centrifugal point around which all later discussions on Croatian orthographic manuals were centred. One was either for or against it (Brozović 2003: 22).

So the actors who planned, i.e. set the norms for, the Serbo-Croatian standard language were grammarians - linguists. Some of these actors may have worked by themselves (individually), but most of them were surely part of a group activity. Their actions were promoted by pressure groups and prestigious members of the speech community (writers or media who chose to follow the prescribed norm). Their actions were supported and promoted institutionally (Matica srpska, Matica hrvatska, The Yugoslav Academy of Sciences and Arts (JAZU)). They were even officially promoted by the state through the educational system, legislation etc., when their orthographic manuals and grammars were decreed as curriculum in the schools.

The initial norms set by Vuk and his followers were based on rural usage, isolated from the surrounding revolutions in technology and science whereas the *elaboration* of it (cf. below) was inspired by the language and needs of an urban society in contact with the trends and achievements of the modern world.

After the establishing of the first Yugoslavia (The Kingdom of Serbs, Croats and Slovenes) in the wake of the First World War, the idea of one nation, one language, and one country led the centralizing forces in the country's capital Belgrade to declare that the official language of the country was Serbo-Croato-Slovenian<sup>121</sup>. The fact that the base dialect of Slovenian, among other things, differs from that of Croatian and Serbian was certainly the reason why no codification (corpus planning) of Serbo-Croato-Slovenian was ever carried out.<sup>122</sup> The prescription of how, first of all, to spell and use punctuation reached the literate public via orthographic manuals, typically consisting of one part with rules on spelling

---

<sup>121</sup> Art. 3 in the 1921 Constitution of The Kingdom Serbs, Croats and Slovenes: „Службени језик Краљевине је српско-хрватски-словеначки.“

<sup>122</sup> This led to paradoxical, but politically correct, titles of the much needed dictionaries between Slovenian and Serbo-Croatian. E.g. Veskić' *Sloveno-Serbo-Croatian dictionary* printed in Belgrade in 1932, where the title indicates a dictionary of either three different languages or just one, while it is in fact a dictionary of Slovenian words translated in to Serbo-Croatian, as it is explained in the one page foreword (Veskić 1932)

and punctuation followed by a dictionary of words considered difficult to spell.

It was through competing orthographic manuals that different views on what the unity of Serbian and Croatian should lead to became evident. The differentiated prescriptions in the different orthographic manuals used in Serbia and Croatia in the 1920's mirrored the fact that the leading Croat linguists, though dedicated Vukovites, did not accept all that was coming from Belgrade. Nevertheless, in 1929 a common orthographic dictionary was produced on the basis of two existing ones, used in Serbia, Montenegro and Bosnia-Herzegovina and Croatia, resp. (see also subsection 2.4.4. below). Before the disintegration of Socialist Yugoslavia the norm-setting bodies of Serbo-Croatian (Serbian and Croatian) were the two cultural institutions Matica Srpska and Matica Hrvatska. However, the attempt to create a common dictionary of Serbo-Croatian through collaboration between the two prestigious quasigovernmental institutions failed and the distinctively Croatian orthographic manual *Hrvatski Pravopis* of 1971, mentioned above, was banned and not published in Croatia until 1990 (Auburger 1999: 282).

Nevertheless, during interwar and post Second World War Yugoslavia, grammars, dictionaries and orthographic manuals adhering to the principle of Serbian and Croatian being one language – Serbo-Croatian were the only officially accepted manuals and as such they represent the corpus planning phases of description and prescription of the period.

During the years following Croatia's independence, under different governments during the 90s and 00s the Croatian language was described and prescribed anew. Linguists in Croatia strove to define the makeup of the Croatian standard and there was a lot of disagreement in very specific issues pertaining to which orthographic rules should govern the standard Croatian. Different orthographic manuals and dictionaries were published both by professional linguists with the backing of prestigious institutions such as *Matica hrvatska* and the nationally dominant publishing house *Školska knjiga* and by linguistic amateurs. These issues were to be solved by the specialists in the area which is why the Council for Standard Croatian Language Norm was established by the Ministry of Education and Sports in 2005. The decisions made in the Council were, however, not always supported by the government. For instance, the decision to standardize the writing of the negated present form of the auxiliary verb in the future tense, Future 1: *htjeti* ('to will') in two words – *ne ću* (which had been in wide use before 1960) instead of

oneword writing – *neću* (which had been the only standardized form since the publishing of the Orthographic Manual of Serbo-Croatian in 1960<sup>123</sup>) was effectively curtailed by the then prime minister Ivo Sanader in December 2005. (Bašić 2012: 62–63).

In 2013 a new Croatian orthographic manual and dictionary (*Hrvatski pravopis*) was published by the state financed Institute for Croatian Language and Linguistics. This *pravopis* was a deliberate attempt, through compromises, to end the controversies between the different views and was ultimately in July 2013 also recommended by the Ministry of Science, Education and Sports for use in primary and secondary schools in Croatia. (Jozić et al. 2013: V)<sup>124</sup>. Some linguists in Croatia stay opposed to the orthographic solutions in this publication and continue to criticize it as for instance Nataša Bašić who does not avoid casting aspersions on the professional knowledge and integrity of the 14 authors and 5 editors of the manual (Bašić 2014).

So, the prescription phase of language planning in Croatia today is still not stable and, when it comes to establishing a prescribed norm of a standard language, discussions may become very heated. Whether the compromise attempted in the 2013 *Pravopis* is met with approval in the general public remains to be seen but as long as it stays the recommended orthographic manual in schools all over Croatia it has a fair chance of success as this medium of *corpus planning* has entered the important prerequisite for its successful implementation which is described as *acquisition planning*, cf. p. 89.

### 2.4.3 Elaboration

Elaboration is a crucial phase when a language is to take on the role as a true standard language as defined by (among others) Josip Silić. A standard language is a polyfunctional institution and as such its function is to be a means of public communication between human beings in all areas of life: science, administration, media and literature (Silić 2006: 10). Elaboration is a constant phase in language planning; the standard language is constantly developing or, more accurately, is constantly being developed so that it may meet the demands to be a means of communi-

---

<sup>123</sup> Published jointly by *Matica hrvatska* and *Matica srpska* (cf. p. 25, above)

<sup>124</sup> The issues of *neću* and *ne ću* and other points of conflicting views are solved by recommending one solution and not prohibiting the other. In the *neću* – *ne ću* instance, the former is recommended and the latter is allowed.

cation under the ever changing social, cultural, communicational and creative circumstances in which its users live (Radovanović 2003: 193; Joseph 1987: 89–90). Radovanović even ascribes the language an ability to constantly adapt to the specific communicative needs within a speech community and the language is as such a mirror image of the societal and cultural structures of speech community (Ibid.).<sup>125</sup> This view is drawn into question by Joseph as he claims that a language as a rule lags behind the ever-changing cultural development, the so-called acculturation (Joseph 1987: 89). Elaborating (acculturating) a language includes the transfer of concepts expressed in other languages and is thus of special interest when dealing with contact-induced linguistic changes.

In the early elaboration of the Neoštokavian standard language, the lexicon needed enrichment, partly because earlier standardisation efforts had centred around other base dialects (Čakavian and Kajkavian dialect in the western parts and slavanoserbski in the eastern part), partly because it was part of nation-building process to develop a language that could meet the communicative need in all spheres of life in an independent state of South Slavs. Zett (1970) goes as far as saying that due to the break with earlier literary tradition, the language corpus planners had to start over: „... brachte der Bruch mit den schriftsprachlichen Traditionen die Notwendigkeit mit sich, auf dem Gebiet des Wortschatzes weitgehend noch einmal von vorn anfangen zu müssen“ (Zett 1970: 16). The elaboration of the lexis happened, mostly through lexical matter replication (lexical borrowing) but also largely through lexical pattern replication (loan translation) both among Serbs and Croats.<sup>126</sup> The main language to be replicated from, the main model language, was German, more so in Croatia than in Serbia. Other languages which also served as models for replication in the elaboration phase were: Czech, Russian, Italian and French (Zett 1970: 16).

In Serbia in the beginning of the 20th century, the language (was) developed explosively due to the fact that prose writers and poets wrote in Serbian and because Serbia was an independent country (recognised in

---

<sup>125</sup> Even though the wording may indicate that the language has a life of its own, it is important to emphasize that language is *not* a biologic creature but a system used and developed by its users.

<sup>126</sup> I will return to the complex matter of transference of lexical, morphological and structural elements from a Model-language to a Replica-language in chapter 3.

1882, de facto so since 1867), where terminology, in Serbian, was needed in all spheres of society. A so-called Belgrade style was developed by the intellectual elite who were in the fortunate position that their variety of Neoštokavian was also the language of administration (P. Ivić 1998: 267). However, this elaboration of Serbian was not followed up by any description or prescription as such. The only 'language planners', i.e. *actors* were the prestigious writers<sup>127</sup> and speakers and they did not care for linguistic norm setting or other grammatical 'trivial matters' (Ibid: 269). The Serbian normative publications<sup>128</sup> still followed the norm and rules set forward by Vuk Stefanović Karadžić during half a century from 1818-1868, 1868 being the year when his language reform gained full legal acceptance. (Brborić 2013: 175).

So, as an excellent example of Ammon's point of language status and function (cf. p. 87), the language usage within a prestigious group in society was turning into usage norm (gaining functional status), but the usage norm was not to constitute the standard language corpus on the basis of which grammarians registered the regularities and prescribed the rules of proper Serbo-Croatian (i.e. the language variety with the officially highest ranking status). Nevertheless, it was within a prestigious usage norm, viz. the result of what was implemented by the prestigious language users, i.e. the writers that elaboration of Serbian happened. According to Haarmann's model (p. 91), this is a case of individual and group cultivating, which may have an impact, but should not have the same impact as an institutionally or by the government promoted planning measure would.

In Croatian, elaboration had taken place before the Vukovites took the lead and used the corpus in the East-Herzegovinian Neoštokavian gathered by Vuk among commoners as the basis for their normative work. Therefore, a large vocabulary existed in Croatian language usage norm which was not admitted into the described and prescribed norm. This vocabulary, naturally, did not go out of usage entirely, but it lost prestige in the Yugoslav context, i.e. it was pushed aside by the Serbian vocabulary in the interwar period and was stigmatized as regionalisms.

---

<sup>127</sup> Weinstein (1983) suggests that these prestigious deliberate language users (writers) be called *language strategists*

<sup>128</sup> e.g. : *Srpski pravopis za školsku upotrebu* by Stevo Čuturilo (1884), *Srpski pravopis za srednje škole* by Milan Petrović (1914) and *Srpska gramatika* by Stojan Novaković (1894)

The fact that, at the time of the unification of Serbian and Croatian, Croatian already had an elaborated vocabulary, distinct from the Serbian ditto, is one of the arguments often used to underline that Serbian and Croatian are too different to function as one standard language. (cf. section 2.3 and subsection 1.1.2 above.)

During the last decade, in Croatia, elaboration of professional terminology has been supported by the government, through the funding of the national term bank *Struna*<sup>129</sup>, established in 2009 with the explicit aim to develop Croatian professional terminology. (IHJJ 2016) *The Council for Standard Croatian Language Norm* initiated the project in December 2006 by suggesting to the Government that such a body should be established. The preceding discussion in the Council and motivation for the project was a growing concern that Croatian professional terminology was replenished and in some areas even substituted with foreign terminology. (VNHSJ 2013: 104) At a subsequent meeting of the Council the chairman, Radoslav Katičić, underlines the importance of the development of a *truly* Croatian terminology with words derived from existing lexical stems in the Croatian dialects or through revivification of older terminology (Ibid: 111). Katičić thus encourages the potential developers of Croatian professional terminology to be loyal to the well-established Croatian puristic tradition, in this instance through *ethnographic*, *reformist* and *archaic* purism (cf. “2.2.3 Purism”). Terminologies for 18 different professional areas have been formally standardized and besides traditional publications of specialized dictionaries, the public can also access the dictionaries through *Hrvatski terminološki portal* (“The Croatian terminology portal”) at <http://nazivlje.hr/>.

In Serbia, in the same period, the equivalent to the Croatian Council, *The Board for Standardisation of the Serbian Language*, has not been successful in applying for funding for such projects. Its attitude towards terminology replicated from other languages is also quite different from the attitude of its Croatian counterpart as is seen in the proposed codex for the Board where the linguist Milan Šipka proposes that a *principle of selective openness* should be applied by normative linguist towards foreign influences and that a *principle of free choice of linguistic resources in individual expression* should be applied while linguistic creativity should be encouraged and made possible within all functional styles in order to harmonize the existing norms within the different functional styles with

---

<sup>129</sup> STRUNA is an acronym for *strukovno nazivlje* – professional terminology

the societal needs of today which he calls the *principle of dynamicity*. (M. Šipka 2006: 297–98). Matić (2012) points out the urgent need for a systematization and pooling of expertise within the field of standardisation of professional terminology and suggest that a law should be passed in parliament dictating that an Institute for the Standardisation of Serbian Terminology be established as an independent institution in Serbia (Matić 2012: 226–27).

#### 2.4.4 Acceptance and Implementation

Which of the changes and developments, implemented by the language users that are to become part of the codified standard depends not only on these changes' adherence to the linguistic system, but more so on its implementation (or lack of same) by its users and/or planners. If the developments (elaboration) are promoted by prestigious groups in society (writers, journalists, politicians, scientists, lawyers) they stand a better chance of being implemented by other users. (cf. Figure 2.4)

Only when accepted and implemented, a language variety or changes within it become a true standard.

Radovanović insists on the division of acceptance and implementation in order to be able to illustrate the fact that official acceptance is not necessarily the same as the implementation. A standard language variety may very well be in use (i.e. implemented) before it is officially accepted, as was the case with Vuk Stefanović Karadžić' language norm which was not officially accepted in Serbia until 1868, but implemented by the then very prestigious group of language users – the writers – much earlier. The opposite may also happen. Linguistic innovations may be officially accepted and promoted by language planners as a part of the standard without being accepted into the usage norm.

The status planning leading to the first *officialisation* or official *acceptance* of the selected language variety in the 19<sup>th</sup> century happened as we have seen, at different times in Serbia and Croatia. As we saw in the in section 1.1 above, different regimes subsequently officialised the Ne-oštokavian language variety either through legislation or through governmental actions as Serbo-Croato-Slovenian (1921) and Serbo-Croatian (1929), Croatian (1941), Serbian (1941)<sup>130</sup>, Serbo-Croatian (1954), Croatian, called Croatian or Serbian (1974), Croatian (1991), Serbian (1992)

---

<sup>130</sup> Not only the fascist government of Croatia and the puppet government in Serbia officialised the language(s) as Serbian and Croatian, but also the Communist Liberation Army (cf. subsection 1.1.5 above)

In Serbia and Croatia and earlier in Yugoslavia the governments conducted language status planning by formalizing the status of language(s) in Constitutions:

Article 138 in the 1974 Constitution of the Socialist Republic of Croatia reads: "*U Socijalističkoj Republici Hrvatskoj u **javnoj** je upotrebi hrvatski književni jezik – standardni oblik narodnog jezika Hrvata i Srba u Hrvatskoj, koji se naziva hrvatski ili srpski.*" (Sofronić 1974: 315) ('In The Socialist Republic of Croatia the Croatian literary language – the standard form of the people's language of Croats and Serbs in Croatia, which is called Croatian or Serbian, is in **public** use.')

Article 146 in the 1974 Constitution of the Socialist Republic of Serbia reads: "*„U Socijalističkoj Republici Srbiji ravnopravni su jezici naroda i narodnosti i njihova pisma“* (Sofronić 1974: 618). ('In the Socialist Republic of Serbia the languages of the nations and minorities and their scripts have equal rights.')

Article 12 in The 1990 Constitution of the Republic of Croatia reads: "*U Republici Hrvatskoj u **službenoj** je uporabi hrvatski jezik i latinično pismo.*" ('In the Republic of Croatia the Croatian language and the Latin script are in **official** use.')

Article 10 in The 2006 Constitution of the Republic of Serbia reads: "*„У Републици Србији у **службеној** употреби су српски језик и ћирилично писмо.*" 'In the Republic of Serbia the Serbian language and the Cyrillic script are in **official** use.'<sup>131</sup>

As we move from the officialisation (acceptance) to the implementation, the agency of the speech community (the recipients of the planning) increases in importance and in both Croatia and Serbia it was of predominant importance that not only the administrators (governments) accepted the proposed planning (status planning) but that the writers (poets and prosaists) accepted and *implemented* the proposed standard (i.e. prestige planning). In Serbia, writers, beginning with the poet Branko Radičević (1824-1853) started writing their works in the language described by Vuk Stefanović Karadžić. (P. Ivić 1998: 211, 236)

In the Croatian lands as well, the Illyrians promoted the Neoštokavian base dialect as a common standard for not only the Croats but also the inhabitants of Montenegro, Bosnia-Herzegovina and Serbia through the literary journal *Danica*<sup>132</sup>, published 1835-1849 thus encouraging

---

<sup>131</sup> Emphasis added in the quoted legal articles.

<sup>132</sup> Also cited on. p. 15 and p.75



writers and poets to write in this particular language variety (Despalatović 1975: 82; Mønnesland 2002: 20).

So, in Haarmann's terms the language planning measures of the 19<sup>th</sup> century amongst Serbs and Croats were first promoted individually by Vuk Stefanović Karadžić, and then promoted by pressure groups consisting of writers through their individual work and through journals (e.g. the *Danica*). In addition the planning measures were institutionally promoted by cultural institutions (e.g. The Yugoslav Academy of Arts and Sciences – JAZU) and officially promoted in virtue of their legal recognition - status. The organisational impact was, applying Haarmann's model, most efficient after the planning measures had been officially promoted by the government.

Hornberger, like Radovanović, singles out the phases of *Acceptance* and *Implementation* using a slightly different terminology as she sees the language's formal role in society as being part of 'language policy planning' whereas the language's functional role in society is part of the 'language cultivation planning'. Despite their different approaches to language planning, Hornberger and Radovanović seem to agree that both *acceptance* and *implementation* are part of extra-linguistic planning (not corpus planning), (cf. Figure 2.5). In Hornberger's model we also saw that in order for a planned measure to be implemented by the language users, it has to reach them through: *a group, education/school, literature, religion, mass media and/or work*. So, in order for the planned language to reach true implementation, it has to be propagated in domains that the users find prestigious for them to be inclined to implement them in their own language use. Haarmann, similarly, points out that the prestige (and thereby the chance of being successfully implemented by the language users) of the planned language changes does not only depend on objective factors such as the administrative (governmental and/or institutionalised) power of the planners but very much so also on the relationship between the planners and the recipients of the planned language. Kaplan and Baldauf contend the same view as they say: "Finally, it is important to remember that proposed solutions must be 'sold' to the population; language change will not necessarily be readily accepted by a population, because language issues are commonly emotion-laden." (Kaplan and Baldauf 1997: 120).

One example of a bad relationship (or bad salesmanship) resulting in unsuccessful planning would be the ministerial decree (of June 1929)

in interwar Yugoslavia to use only one orthographic dictionary throughout Serbia, Croatia, Bosnia-Herzegovina and Montenegro. The Ministry of Education in interwar Yugoslavia established a committee with the distinct task to root out the differences between two competing orthographic manuals *Pravopis hrvatskoga ili srpskoga jezika* (Boranić 1921) and *Pravopis srpskohrvatskog književnog jezika* (Belić 1923) which were used in Croatia and in Serbia, Montenegro and Bosnia-Herzegovina. The preliminary work on this unified orthographic manual was done by the author of the Serbian one, Belić and the result of the unification was, not surprisingly, a slightly modified version of his original work. This orthographic manual and dictionary, which is commonly known as the ‘agreed’ orthographic manual (*dogovorni pravopis*) was published under the names ‘Orthographic manual for all primary, secondary and vocational schools in the Kingdom of Serbs, Croats and Slovenes’ (*Pravopisno uputstvo za sve osnovne, srednje i stručne škole Kraljevine Srba, Hrvata i Slovenaca*) and ‘Orthographic manual for all primary, secondary and vocational schools in the Kingdom of Yugoslavia’ (*Pravopisno uputstvo za sve osnovne, srednje i stručne škole u Kraljevini Jugoslaviji*) (Samardžija 2008: 18–19; Brozović 1998: 15). Other publications concerning so-called common professional terminology followed, but as Brozović points out ‘common’ equalled ‘Serbian’ and furthermore, the unified (‘common’) orthographic manual and terminologies were *rejected* by all in Croatia who were not a part of the regime, i.e. by writers, non-governmental publishing houses, the press (media), journals, etc. (Ibid:)

So, the lack of prestigious implementers (i.e. their lack of popularity) can seriously reduce the efficacy of planned language changes and if a change is not propagated in the domains outlined by Hornberger it will not be implemented.

The significance of media in setting the standard of what is proper and prestigious language usage is commonly recognised as early as in 1933 (Joseph 2006: 117) and was in the Serbian, Croatian context pointed out already in 1958 by Ljudevit Jonke, one of the signatories of the 1954 Novi Sad Agreement in his criticism of how the Serbian Ekavian variant was the only variant in use in the federal radio news emitted to the whole of Yugoslavia (SFRY). This constituted a violation of the 4th point of the Novi Sad Agreement pertaining to ‘the absolute equality of the two pronunciations Ijekavian and Ekavian’ (Jonke 1971: 288).

In what measure the prescribed norm of today’s linguistic authorities in Serbia and Croatia is implemented by the language users, we will

return to in the analysis of the empirical data, but it is clear that the language of the media, and the people present in the media have an even larger influence today than 70 years ago on what is implemented and subsequently *expanded* and what is not.

#### 2.4.5 Expansion and Cultivation

In order for the standard variety to become a true standard language of a society its use has to expand not only into the different functional styles (registers) but also geographically so. It has to take over the communicative and creative function in a wider geographical area and penetrate all social layers and by doing so it suppresses the language varieties in these areas and local communities (Radovanović 2003: 194).

This particular phase in Radovanović's model seems to be somewhat at odds with the otherwise commonly accepted view that the standard variety/language is a supraregional language variety as defined by Southerland and Katamba in: O'Grady et al. (eds.) *Contemporary Linguistics – An introduction* "This superposed variety is employed by the government and communications media, used and taught in educational institutions, and is the main or only written language." (Southerland and Katamba 1997: 541).<sup>133</sup>

However, if we understand Radovanović's description of expansion as an illustration of how the standard variety becomes the only variety in which it is accepted to write and furthermore that the standard variety is the variety used when communicating interregionally (allowing the regional variety to be used interregionally), his view does not differ essentially from the common view.

Furthermore, the expansion is linked to the cultivation of the standard which is done through teaching, maintaining and propagating it through the schools, publishing houses, media etc. (Radovanović 2003: 194). In other words, through usage of the standard by prestigious actors (schools, media, and publishing houses) its usage is also expanded into geographical areas and areas of society where another language variety (or language) has hitherto been the main idiom of communication.

The term 'cultivation' does not carry the same meaning in Radovanović's model as it does in the cited works of Haugen, Kloss, Ammon, Hornberger, Haarmann and others. Whereas it is otherwise a term covering the modernisation of an idiom (making it cultivated), Radovanović

---

<sup>133</sup> Also quoted in section 2.1 above.

uses the term to cover the deliberate spread of its usage, i.e. cultivation of the language users, not the language. The more common (and broader) definition of language cultivation seems to coincide with Radovanović's *elaboration*.

On the other hand, the activities involved in Radovanović' definition of cultivation (propagating, teaching and maintaining the standard) coincide, at least partly, with what is covered by the term *Acquisition planning*. In schools children are taught how to pronounce, read and write on the basis of manuals accepted by the ministries of education. Finally, the media also play a role in mediating the rules of the language planners in as much as the media houses in their employ have language editors who are in charge of editing texts before they are published or aired. In the early 21st century Croatia, according to Kordić (2010), the diligence put into correcting and purifying the Croatian language rubbed off on almost everybody. You risked being corrected by your own butcher, were you to use an expression which was deemed un-Croatian (Kordić 2010: 40). Expansion thus naturally follows Implementation, as speakers need first implement an innovation in their own speech before they begin to contribute to the expansion of the use of this innovation simply by being heard by other speakers.

In the latter instance it is not an issue of planned *expansion*, but rather a question how good the planners are at influencing the speech community, i.e. motivating speakers first to *acquire*, then to *implement* and thus *expand* the use of the planned changes.

#### **2.4.6 Evaluation and Reconstruction**

Having achieved both the status and the function of a standard language, the norm will be under continuous evaluation which will show in which areas it lacks communicative power. The language users will find a way to communicate despite these shortcomings, but they will have to use linguistic means (lexical as well as grammatical means) which are not part of the standardised norm. This is where linguistic innovations are made by the users. Whether these innovations are spread to other users and accepted into the usage norm, and whether these innovations are to be a part of the standard language depends on whether there is 'a confirmed need to change the norm' (Radovanović 2003: 194). If this is the case the tenth phase, the reconstruction phase commences. In the reconstruction phase, phase 2, 3 and/or 4 (description, prescription and/or elaboration) are in play and thus the process is 'restarted'. This cyclic process of ten phases or procedures through which an innovation must

go in order to become the norm is according to Radovanović a model via which language planning should be explained<sup>134</sup>. This means that the proposed model is not only a means to explain the emergence of a standard language, but also a way in which to view all changes of the standard language norm (Radovanović 2003: 195; 2000: 31; 1996: 15).

Following the wars in the 1990s, in both Serbia (in 1997) and Croatia (in 1998 and again in 2005) boards of scholars (which have been referred to above) were appointed to maintain and take care of Croatian and Serbian. In Serbia, *The Board for Standardisation of the Serbian Language* was established by academies of sciences and art, universities and leading cultural institutions in Serbia, Montenegro and Republika Srpska. In Croatia, in 2005 *The Council for Standard Croatian Language Norm* was established by the Ministry for Education, Science and Sports and the same ministry appointed all its members. It was dispersed in 2012. Earlier, in 1998 another similar council had been established by the Ministry of Education and Technology only to be dispersed in 2000.

The objectives and tasks of the two bodies differ somewhat, possibly mainly due to their different status (in Croatia governmental, in Serbia non-/quasigovernmental) but both bodies have as their main objective to set the language norm: The Croatian Council was to: 'pass decisions in connection with the further *setting of language norms* for the Croatian standard language'<sup>135</sup>, The Serbian Board took upon itself the task of: 'systematically *setting the norms* for the Serbian language, including both Ekavian and Ijekavian pronunciation, in general and in detail, and the production of matching documents and manuals as well as passing acts which should secure the viability of authoritative innovations within both norm setting and usage of language.'<sup>136</sup> (Emphasis added)

These two bodies should not, however, be seen as the sole perpetrators of language planning. Other institutions in both Serbia and Croatia have been involved in corpus planning. In Croatia: the Croatian Academy of Sciences and Arts (HAZU) and Institute for the Croatian Language and

---

<sup>134</sup> "Ukoliko[...] za promenu norme postoji utvrđena potreba..."

<sup>135</sup> Original: „...donijeti rješenja u svezi s daljnjim normiranjem hrvatskoga standardnog jezika“

<sup>136</sup> Original: „систематско нормирање српског језика, с екавским и ијекавским изговором, свеобухватно и у појединостима, и израда одговарајућих докумената и приручника, као и доношење аката који би обезбеђивали проходност“  
меродавних иновација у нормативистици и језичкој пракси (kilde: Rastko<sup>1</sup>)

Linguistics (IHJJ)<sup>137</sup> (Langston and Peti-Stantić 2011: 344; 2014: 147). With regards to Serbian, several institutions have been and are involved in language planning via The Board for Standardisation of the Serbian Language as they, not the government, founded this board (Projekat Rastko 2015).

So, we may conclude that the actors in both Serbia and Croatia who *plan language* are to be found in all of the five types that Kaplan and Baldauf as well as Haarmann outline (1) Government agencies, (2) Education agencies, (3) non-/quasigovernment agencies and (4) other organisations as well as by (5) individuals. There were and still are an abundance of language advisers, who put forward their advice in not only scientific periodicals such as the Croatian *Jezik* ('Language') and the Serbian *Naš jezik* ('Our language') but also in newspapers, magazines and in the electronic media there are programs and columns dedicated to linguistic questions.

## 2.5 Language planning and Contact linguistics

Research into language planning, language policy and the politics of language is, in itself, a most interesting field of study and I believe that it has, together with a discussion of the somewhat complex definitions of the standard language and an outline of the strategies of applied linguistic purism, provided the tools needed to discuss, systemize and elaborate on what has been done to and for the standard languages Serbian, Croatian and Serbo-Croatian during the last approx. 180 years

In the next chapter I will turn my attention to explanations of how changes come about in languages, more specifically to how linguistic innovations may be triggered by contact. Explanatory frameworks and theories which explain this phenomenon as well as methods to analyse it are found within the scientific fields of contact linguistics and historical linguistics.

The reactions of language planners to so-called contact-induced language change and not least the impact of language planning measures aiming at diminishing the influence of contact is where the scientific fields of contact linguistics and sociolinguistics meet. The historical overview and the explanatory frameworks of language planning presented, discussed and exemplified in the present chapter will serve as a

---

<sup>137</sup> Hrvatska akademija znanosti i umjetnosti (HAZU) and Institut za hrvatski jezik i jezikoslovlje (IHJJ)

necessary context in which specific language changes, linguistically explained by means of universal theories and models of language change observed in Serbian and Croatian modern language usage, are analysed.

### **3. Language Change and Contact**

The following chapter is devoted to the field of language change, especially language changes induced by contact with other languages. Language change is traditionally studied in the field of historical linguistics where changes in linguistic structures that have happened through a period of several hundred years of a particular language or languages are studied. The motivation for the changes (the human factor) is usually not the focus of language change studies. By focusing on contact with specific languages, and on a period of great societal changes I will apply the theoretical findings from scholars in historical linguistics by drawing upon a general theory of language change by Henning Andersen (2001; 2008), and with the aid of other insights provided by Bernd Heine and Tania Kuteva (2005, 2006, 2010), Hopper and Traugott (2003) as well as Brinton and Traugott (2005) discuss the phenomena of grammatical and lexical change (3.2). Language change is, however, also viewed as on-going process (a gradual change), so subsequently I will add the insights into gradual contact-induced language change (3.3) offered by Sarah Thomason (2001, 2010), Bernd Heine and Tania Kuteva (2005, 2006, 2010) Henrik Gottlieb (2006) and Yaron Matras (2009, 2010).

In principle, every time a linguistic innovation is made in a language, the inventory of said language is changed. The replication of lexical matter (loanwords) does not alter the structure of the language but it certainly enriches it. When word-formational patterns are replicated, structural changes within word-formation in the receiving language may occur and if a grammatical category is replicated, the structure of the receiving language is likely to change. A replication of some linguistic matter (lexical or non-lexical) may in time instigate changes in the linguistic pattern (the structure) of the receiving language which is also called a grammatical change.

#### **3.1 Perspectives on change**

Change may be viewed as an instantaneous event as for example the legal inauguration of Croatian as the official language in Croatia which happened at the moment when the Croatian constitution in 1990 was passed. However, change is very often not instantaneous but a gradual process for a shorter or longer period of time which may be repeated at different points in time, as e.g. the changes in Croatian standardised norm due to the puristic tendencies in Croatian standardisation efforts in the late 19<sup>th</sup>



century and the late 20<sup>th</sup> century or due to the efforts towards unification of Serbian and Croatian in the mid 19<sup>th</sup> century, the interwar period and in the years following the Second World War.

Contact-induced language change may come about very quickly, as when Croat speakers altered their language usage in order to comply with the new political reality in the 1990s by increasing their usage of domestic linguistic matter instead of replicated matter or when a new phenomenon needs a name which is typically replicated from another language (*kompjutor* < 'computer') and only subsequently this replication is substituted by a new domestic word (*računalo* – 'computer'). The replicated *kompjutor* is today recognised as a word in usage in Croatian but are by normativists language planners relegated to be used in conversational functional style and the jargon of IT-specialists.

Other linguistic changes, structural changes, enter the usage norm in a much slower fashion as e.g. the change in word-formation of compounds which has led to nominal compound formations as e.g. *menadžment sposobnosti* ('management skills'), *nana čaj* ('mint tea'), *internet veza* ('Internet connection'), which have yet to enter the prescribed standardised norms of both Serbian and Croatian, to increased word-formation of nominal semiword-compounds with semiwords such as e.g. *etno-*, *jugo-*, *-log*, *-logija* and *-erija* in e.g. *etnoselo* (ethnic village), *jugonostalgija* (Yugo-nostalgia), *ćutolog*, ('silent, introvert person'), *papirologija* ('excessive amounts of paper', 'red tape') and *piterija*, ('pita place') which have entered the usage norm but only in certain instances also the standardised prescribed norm, or to the increased article-like usage of *jedan* ('one') and *neki* ('some'), interpreted by some as a consequence of contact with article-languages, as in e.g. :

*Povremeno bi mu i neki prijatelj donio litru domaće rakije.*

('From time to time, a friend brought him a litre of moonshine')

*Nekoliko dana prije toga od jednog prijatelja u Munchenu čuo sam kako...*

('A couple of days earlier, I heard from a friend in Munich that..')

### 3.2 Language change

Coseriu, whose insights contributed to the definition of a standard language (see section 2.1.1 above), approaches language change through an "integrated functional-pragmatic theory of language and language change" (Nedergaard Thomsen 2006: 313). Coseriu distinguishes between three ways of approaching the problem of language change. Firstly, the problem of the rationale, i.e.: Why do languages change? Why do they

not stay the same? Secondly, the general problem of language change, which has to do with the conditions under which language changes occur. Thirdly, the historical problem of a specific change, i.e. investigations of gradual change of a specific linguistic feature or element, typically, over longer periods of time. (Coseriu 1974: 56).

Coseriu's general problem is connected to one of the research questions in this study, viz. "What sort of impact does the language planning and language policy of linguistic authorities concerning contact-induced linguistic changes have on language usage?" as this question pertains to the conditions under which the investigated linguistic innovations may or may not be accepted into the usage norm.

The general problem of language change, the conditions under which language changes occur may, especially when considering (national) standard languages, be immensely political. Hence it is appropriate and fruitful to describe and analyse these changes by applying theoretical language planning frameworks as outlined in the preceding chapter. The general problem of language change is, as is the problem of rationale, linked to the communicative needs of its speakers, but also to the actual setting in which the language-changing speakers find themselves, to the speakers' attitudes toward innovations, to the symbolic import of the innovation, to the inspirational source of the innovation, etc.

Coseriu's third question or problem, which, among other things, concerns the specific gradual linguistic changes, is connected to another of the research questions of this study, viz. "Do any of the linguistic innovations constitute a change of the languages themselves?" and in the present chapter views of what the necessary conditions are for regarding a linguistic innovation as a linguistic change and what lexical and grammatical change comprises will be presented and discussed. In the following chapter, on the specific linguistic innovations, the selected types of linguistic innovations will be exemplified and it will be established whether they constitute language change.

The answer to Coseriu's first question "Why do languages change?" is closely connected to why languages exist. To this question I will adhere to the functionally sound assumption that the primary *raison d'être* of language is communicative. In order for a language to function as a means of communication, a language needs to be systematic. In other words, a language 'works' if it is ruled by structural regularities, so that a language user (a speaker) may produce speech and understand speech produced by other speakers.

Language is in constant use and is continuously adapted by its users to accommodate their communicative needs. This adaptation may lead to language change, or rather: make the language users (the speakers) innovate and accept innovations in their language, which if accepted and implemented by a sufficient amount of speakers, will become norm, thereby constituting an innovation in their common language usage. Because speakers of one language are not immune to influences from other languages, the innovations and ultimately the language innovations that they make, may very well be motivated by contact with one or several other languages.

The historical problem, the investigation and study of the individual language changes, is a problem that is investigated in the field of historical linguistics, where researchers investigate and describe the origin of individual changes and where explanations (theories) on language change are developed. One of the theories has been developed by Henning Andersen who explains the origins of a change in the language used by a whole speech community in the following way:

A change originates in one or more speakers who make an innovation, which is then gradually accepted and put into usage by other speakers and finally generalized in the speech community. (Andersen 2008: 21). Andersen envisages the process like this:

The normal course of events through which a new expression originates, gains currency, and becomes established as part of a tradition of speaking is the following: one or more speakers (i) make a (primary) innovation and (ii) actualize it in usage; other speakers (iii) adopt the new expression and (iv) actualize it in their usage; if the new expression is used widely and long enough, new cohorts of speakers (v) will acquire it as an integral part of their competence, and (vi) actualize it in their usage; the new expression becomes generalized in the community through repeated cycles of (iii)-(vi). One can speak of such a series of overlapping kinds of innovation as a 'change scenario' and of the constituent innovation types (i)-(vi) as 'subchanges'. (Andersen 2008: 21–22)

It follows that far from all innovations reach the point where they are generalized in the speech community and therefore never become part of the norm. As mentioned above, the usage of a given language is constantly being innovated by its users. Some of these innovations stay on

and become a part of the language use of so many language users that they become part of the usage norm, i.e. a usage which many language users implement in their own usage. Innovations which become part of the *usage norm* constitute language changes. Such changes may be accepted by language authorities or not, i.e. they may be allowed to enter the standardised norm of the language, the standard language, or not. The potentially demotivating factors which curtails the generalisation "in the [speech] community through repeated cycles (iii)-(vi)" is not the focus of Andersen's envisaged "normal course of events through which a new expression originates, gains currency, and becomes established as part of a tradition of speaking" as the focus of language change studies is rather to trace the origin of a change that can be observed by diachronic comparison of language usage at two or more, often very distant, points in time. Still, the insights of Andersen and other historical linguists provide us with a notion of the language internal mechanisms which lead to linguistic innovations and give us a general idea of which path an innovation follows in order to (a) be implemented by speakers and (b) induce change of the pattern in a language.

The historical linguist Andersen's description of the change scenario has a striking resemblance to the description provided by sociolinguists Radovanović, Haarmann and Hornberger concerning two of the steps that a language goes through to become a true standard language, namely the sixth and seventh step in Radovanović's model: *implementation* and *expansion*. This should, however, not come as surprise as we are in both instances dealing with changes or innovations which need to be implemented by enough users in order for the change to become part of the norm. The differences between the two perspectives are (at least) twofold: 1) Language planning measures constitute a top-down process whereas Andersen's innovations are a bottom-up process. 2) We are, therefore, dealing with different agents or actors with different or even opposing motives; a language planner seeks to achieve uniformity while the individual language innovator may invent his innovations to achieve uniqueness.

A linguistic innovation may consist of new lexemes, new semantic content of existing features, new topology, new word-formational patterns, new morphosyntactic structures, new phonetic phenomena. The acceptance and implementation of these innovations then constitute a variety of language changes. The changes may then be labelled lexical,

phonological, phonetic, morphological (in inflection, derivation and compounding), syntactic, stylistic or pragmatic (Görlach 2004: 3–4). The innovations I have chosen to investigate in this dissertation fall within the lexical (replicated nouns), morphological (viz. word-formation through nominal compounding and derivation), and syntactic innovations (marking of indefiniteness).

A language change need not manifest itself as a new occurrence; an increase in the usage of certain features may also be considered a language change. Heine and Kuteva (2005) call this *a rise from a minor to a major use pattern* and argue that the increased usage of an existing pattern, e.g. word-formation by compounding, may lead to a language change. (Heine and Kuteva 2005: 44–45). In these instances language change does not presuppose innovations as easily discernible as when a new lexeme is either replicated from another language or created to denote a new phenomenon<sup>138</sup> or when pronunciation of a vocoid changes through time<sup>139</sup>. Heine and Kuteva’s approach is extremely useful when the research is concerned with on-going, gradual changes in the structural make-up of a language (as in the present study) and seeks to set up the predictable outcomes of these tendencies on the basis of already proved structural changes in other languages.

In the various works on language change a line between what is lexical and what is grammatical is almost always present. Some scholars only deal with grammatical change; others focus on the creation as well as occurrence of new lexemes and the adoption as well as adaption of lexemes from other languages. As this study deals with both kinds of innovations, the following subsection on grammatical and lexical change will serve as a guide to understanding the different mechanisms which have been observed in both types of linguistic innovations.

### 3.2.1 Grammatical change and lexical change

I find that, in order to enter into the following discussion of grammatical and lexical change and subsequently the terms grammaticalisation and word-formation, it is crucial to begin with a clear definition of what is considered *content* and what *function*.

---

<sup>138</sup> E.g. We suppose that the lexeme “aeroplane” did not exist before the invention of this type of machinery

<sup>139</sup> E.g. Old English [a:] turning into Modern English [ou] via Middle English [ɔ:]. [ba:t] > [bɔ:t] > [bout] - ‘boat’ (Murray 1997: 314).

*Content items* (also called content words, lexical items, lexical words) are defined as lexemes that are used to describe things, actions and qualities (traditionally ascribed to the word classes: noun, verb and adjective). Content items refer to something or someone extra-textual, i.e. they possess referentiality to imagined, believed or real world units (animate as well as inanimate), to activities and situations as well as inferential qualities (such as good, bad, green, ugly, modest, etc.) Content items belong in the open, productive lexical classes, i.e. new content items are created all the time and there are many of them.

*Function items* (also called function words, grammatical words, grammatical items, grammatical morphemes) are used to signal relations intra-textually (ascribed to the word classes: adverb, preposition, pronoun, conjunction, and auxiliaries as well as bound morphemes such as derivational and inflectional affixes). Function items belong to the closed, unproductive classes, i.e. new function items are rarely created and they are small in numbers and sometimes they are non-referential.<sup>140</sup>

### ***Grammatical change***

In this study the term *grammatical change* will refer to a change in the grammatical system or grammar of a language. Grammatical change will also refer to the change in the possible grammatical functions of a linguistic item (a lexeme, a morpheme). A change of function from a content item to a function item is referred to as grammaticalisation. Much research has been done in this field, where the main objective has been to ascertain the pathways of change, historically within one language (historical linguistics) and the systematicity of these changes across languages (typology). In historical linguistics, emphasis has been put on *grammaticalisation*. Following the definition in Hopper and Traugott (2003), grammaticalisation covers the process of a “subset of linguistic changes through which a lexical item in certain uses becomes a grammatical item, or through which a grammatical item becomes more grammatical” (Hopper and Traugott 2003: 2) or in Christian Lehmann’s terms: “grammaticalization is a process which turns lexemes into grammatical formatives and makes grammatical formatives still more grammatical” (Lehmann 1985: 303). Finally, we may look upon what has to have happened in order for a grammaticalisation to take place. In the following

---

<sup>140</sup> For a discussion of different labelling and understanding of content and function items, see Brinton and Traugott (2005) pp. 9-18.

list of parameters, presented here in the wording of Heine and Kuteva (2005), it is only the first parameter (a.) *extension* that is a pragmatic prerequisite for grammaticalisation. The remaining parameters are consequences of *extension* and their order could be seen as the 'normal' order of changes.

*Parameters of grammaticalisation:*

- a) extension, i.e. the rise of novel grammatical meanings when linguistic expressions are extended to new contexts (context-induced reinterpretation)
  - b) desemanticizing (or "semantic bleaching"), i.e. loss (or generalization) in meaning content
  - c) decategorialization, i.e. loss in morphosyntactic properties characteristic of lexical or other less grammaticalised forms, and
  - d) erosion (or "phonetic reduction"), i.e. loss in phonetic substance
- (Heine and Kuteva 2005: 15)

*Clines of grammaticality*

This 'normal' order is also reflected in the traditional way of regarding grammaticalisation in which grammaticalisation is viewed as a cline of grammaticality which a lexeme follows on its path from being less to more grammatical.

*Cline a*

content item > grammatical word > clitic > inflectional affix  
(Hopper and Traugott 2003: 7) (Nørgård-Sørensen, Heltoft, and Schøsler 2011: 11)

The grammaticalisation is a process which results in the ascription of a new function to an existing content or function item and could be described as gradually going from having no grammatical function to having one (or several) as illustrated by van der Auwera in the following cline.

*Cline b*

no function > lower degree of function > higher degree of function  
(van der Auwera 2002: 21)

A typical example of a grammaticalisation which follows clines a and b would be the grammaticalisation of the prehistoric Common Slavonic lexical verb: *jesmĭ* ('(I) am') to the auxiliary verb *jesmĭ* in *pisa-l-a jesmĭ* ('(I) have written') in Old Slavonic to the Polish inflectional affix: *-m* in *pisa-t-a-m* ('(I) wrote'):

- a) content item > grammatical word (> clitic) > inflectional affix  
 b) no function > lower degree of function > higher degree of function
- |                     |          |   |                                     |          |   |                          |
|---------------------|----------|---|-------------------------------------|----------|---|--------------------------|
| <u><i>jesmĭ</i></u> | ('I) am' | > | <i>pisa-l-a</i> <u><i>jesmĭ</i></u> | ('I) am' | > | <i>pisa-t-a-<u>m</u></i> |
| VB                  |          |   |                                     | AUX      |   | AFFIX                    |

(Nørgård-Sørensen, Heltoft, and Schøsler 2011: 9)

This gradual change of an item from being content (lexeme) to function (affix) has also been illustrated by Hopper and Traugott as that of a concept being one of content to one that is increasingly functional (Field 2002: 32).

a basket *full* (of eggs...) > a cup*ful* (of water) > hope*ful*  
 (Hopper and Traugott 2003: 7)

However, and most importantly, the cline of grammaticality does not cover all grammatical changes in as much as loss of grammaticality (degrammaticalisation) is also a grammatical change and in as much as changes in word-formational patterns or paradigmatic relationships are also considered changes in the grammar.

### ***Lexical change***

The term *lexical change* will in this study refer to the implementation of new lexemes and morphemes and the reimplementation of old lexemes and morphemes. Lexical change will also refer to new semantic content in existing words and morphemes as well as increased content in hitherto function items. The latter change is termed *lexicalisation* and is understood as the process of a historical change that results in new lexical/content items (Brinton and Traugott 2005: 96–98). Lexicalisation may thus also be viewed as the reverse of grammaticalisation also called *degrammaticalisation* (van der Auwera 2002: 20).

Andersen (2008: 9) uses the terms relexicalisation and delexicalisation to cover the semantic change which leads to either new and added



content – relexicalisation, e.g. *mouse* - RODENT > *mouse* – RODENT, TIMID PERSON, BLACK EYE, HAND OPERATED DEVICE USED FOR COMPUTER INPUT<sup>141</sup> or to cover the semantic bleaching (delexicalisation) of a content items which typically happens concurrently with its grammaticalisation.

### *Clines of lexicality*

As opposed or as an addition to the cline of grammaticality, a cline of lexicality would illustrate the path of linguistic items in inferential, syntactic or morphological linkage with different referents (content) or combinations of referential and non-referential items (content and function items) into one linguistic item, one lexeme following a path from analytical phrases over complex items/lexemes into simple lexemes.

### *Cline c*

phrasal > complex > simplex

An English example would be the gradual development of the lexeme 'gent'. At the first stage we have a typical English NP, consisting of an adjective 'gentle' and a noun 'man'. At the first stage it is possible to insert an adjective between 'gentle' and 'man' thus confirming that we are dealing with two distinct lexemes (gentle and man), e.g. (a) 'gentle, sweet man'. At the second stage this possibility falls away and we have a complex lexeme, consisting of two, still individually analysable elements (morphemes), but with a common referent. At the last stage this complexity has fallen away and we are left with a simplex lexeme.

(a) gentle man > (a) gentleman > (a) gent

When lexicalisation is viewed as the reverse of grammaticalisation (degrammaticalisation), it refers to the situations when a function item takes on content, or when a quality assigning content item (an adjective) takes on the content of a noun, i.e. takes on referential meaning as exemplified below with *ism* and *elderly*

---

<sup>141</sup> The definitions are abbreviated versions of definitions retrieved from the on-line Merriam-Webster Dictionary (<http://www.merriam-webster.com>)

*Cline d*

(bound)function item > depend. content item > independ./free content item

-ism

>

ism

(an) elderly (man) > (the) elderly

This analysis may be transferred to examples of bound lexical morphemes in replicated compounds in Serbian and Croatian such as *strit-art* ('street art'), *art-boks* ('art boxing') or *pop-art* ('pop art'), where it has been observed that *art* has gradually become a full, free morpheme, a lexeme. (Slijepčević 2013: 328). The morpheme 'art' was not viewed as even a bound morpheme when it was originally replicated into Serbian and Croatian because replicated compounds, generally, are not replicated as compounds but as simple lexemes and therefore cannot be analysed as compounds in the replica language unless its individual components have already entered the replica language (Ibid: 327).

Undoubtedly it is so that, with the increase of replicated lexical matter, where one part is made up by a recognisable string of phonemes, this string of phonemes may be reanalysed as an independent but bound morpheme and this bound morpheme may then be reanalysed as a free morpheme, i.e. a lexeme in the same sense as the derivational affix '-ism' in English is reanalysed as the lexeme (noun) 'ism'.<sup>142</sup>

Other, older examples of compounds in the model language which are, at least initially, analysed as a simple lexeme in the replica language would be *džentlmen*, *pejperbek*, *biftek* - replicas of 'gentleman', 'paperback' and 'beefsteak', all replicated from an English model, though in English they are perceived as complex lexemes, i.e. compounds.<sup>143</sup> Following Brinton and Traugott (2005), I would regard the lexicalisation of the bound lexical morpheme *art* as an instance of *Lexicalisation as increase in autonomy* which in the literature is often exemplified by the declitisation of clitics. Clitics are bound morphemes and always a phonological part of another word-form. (Brinton and Traugott 2005: 57–59). The clitics that go through declitisation to acquire an autonomous use, as do the bound morphemes which are separated from the items to which they were bound, like '-ism' or *art*. The autonomous use of the bound lexical morpheme "unbinds" the lexical morpheme, thereby turning it

---

<sup>142</sup> E.g. "When you open a textbook on contemporary political philosophy, you will see a big pile of these "-isms"" (Liu 2015: 101)

<sup>143</sup> the examples, but not the analysis, were collected from Filipović (1990)

into a free lexical morpheme, viz. a *lexeme*. I assume that the path or cline which the bound lexical morpheme must climb is one where it is originally only combined with very few other lexical items (here exemplified by *pop*), then it is combined with an increasing of lexical morphemes (shown here as X) before it is used as an independent lexeme:

*The extension of 'art'*

string of phonemes	> dependent item	> free content item
-content	> +content, -autonomy	> +content, +autonomy
pop-art	> art-X/X-art	> art

I will in my later discussion of such wordforms use the term *semiword* to denote the items that are at this intermediate stage, where they may be attached to many different lexemes, but stay bound in the sense that they are not independent lexemes, even though they carry lexical meaning, thereby being a content item.

Brinton and Traugott (2005), however, do not consider word-formation to be lexicalisation and thus object to the labelling of any word-formation such as *clipping* (as in the “ism”-example) as lexicalisation, neither do they consider “simple borrowing without formal or semantic change” to be lexicalisation (Brinton and Traugott 2005: 96–98).

Van der Auwera, on the other hand, defines lexicalisation as “the making of a lexical item out of something other than a lexical item” (van der Auwera 2002: 21) Besides in the instances which subsume grammaticalisation, he also includes compounds such as *songwriter* because even though both *song* and *writer* are lexical items (lexemes) on their own, *songwriter* only takes on the meaning WRITER OF SONGS after it has been *lexicalised* (Ibid: 26). In other words, lexicalisation may also be viewed as equal to word-formation of content words.

Finally, lexicalisation may be viewed as the formation of all words. Word-formation has been described extensively, more or less as long as languages have been described. The ways in which words are formed are described, mostly with the diachronic assumption that one lexeme is derived from another or as the result of a combination of two or more other lexemes. The lists of types of word-formation vary in diversity but the main corps consists of the following (exemplified with English (a1-7) as well as Serbian or Croatian examples (b1-7)):

<u>derivation</u>		
a1)	<i>public + ity</i>	> <i>publicity</i>
	A+ affix	> N
b1a)	<i>izlet+ište</i>	> <i>izletište</i>
	N+affix	> N
	(excursion+affix)	> ('vacation area')
b1b)	<i>papir + o + logija</i>	> <i>papirologija</i>
	N+SW <sup>144</sup>	> [N+interfix+SW] <sub>N</sub>
	paper + SW[knowledge of]	> ('excessive amounts of paper')

<u>conversion</u>		
a2)	<i>(a) finger</i>	> <i>(to) finger</i>
	N	> VERB
b2a)	<i>Teget</i> <sup>145</sup>	> <i>teget</i>
	N	> A
	('Teget')	> ('navy blue')
b2b)	<i>jedan</i>	> <i>jedan</i>
	NUM	ART
	('one')	('a/an')

<u>clipping</u>		
a3)	<i>hamburger</i>	> <i>burger</i>
	N	> [clipped N] <sub>N</sub>
b3)	<i>električni napon</i>	> <i>elektronapon</i>
	[A+N] <sub>NP</sub>	> [clipped A+N] <sub>N</sub>
	('electrical current')	> ('electrical current')

<u>blending</u>		
a4)	<i>breakfast, lunch</i>	> <i>brunch</i>
	N , N	> [clipped N + clipped N] <sub>N</sub>
b4a)	<i>putopis, bioskop</i>	> <i>putoskop</i>
	N N	> [clipped N + clipped N] <sub>N</sub>
	('travelogue'), ('cinema')	> ('pictorial travelogue')

---

<sup>144</sup> SW is an abbreviation for *semiword*, which is a label which I, following Bauer (1998) give morphemes that in function are akin to affixes but in content to lexemes.  
<sup>145</sup> The proper noun *Teget* was an abbreviated form of the surname 'Tegethoff', a 19th century Austrian Admiral

<u>compounding</u>			
a5)	<i>gentle man</i>	>	<i>gentleman</i>
	[A+N] <sub>NP</sub>	>	[A+N] <sub>N</sub>
b5)	<i>djel- -o- krug</i>	>	<i>djelokrug</i>
	root interfix N	>	[root+interfix+N] <sub>N</sub>
	(‘act’ ‘-’ ‘circle’)	>	(‘sphere of activity’)

<u>acronyms</u>			
a6)	<i>N A T O</i> [‘ɛn ‘eɪ ‘ti ‘əʊ ]	>	<i>NATO</i> [‘neɪtəʊ ]
	letters pronounced	>	N
b6)	<i>S I D A</i> [‘sə ‘i: ‘də ‘ɑ:]	>	<i>SIDA</i> [‘si:da]
	letters pronounced	>	N
			(‘AIDS’)

Finally, some words are coined or created without combining other existing forms. This process is labelled coinage or root creation and can be exemplified in English by the onomatopoeia (a7a) *cock-a-doodle-do* and the noun (a7b) *hobbit* and by the Serbian and Croatian version of cock-a-doodle-do: (b7) *kukuriku*

With the exception of conversion (a2 and b2) and coinage (a7 and b7), the types of word-formation listed here all involve combinations of existing linguistic items. As we will see in the sections in chapter 4 on nominal compounds, semiwords and affixes, examples of clipping and blending (as in a3, a4 and b3, b4), i.e. the possibility to combine a certain string of phonemes with other lexemes is often preconditioned by a lexicalisation, i.e. an ascription of content to a hitherto unrecognised item, as e.g. *-skop*, a string of phonemes first replicated as part of nouns (e.g. *bioskop*), then at some point recognised as having a particular content in the replica language and subsequently gaining the status and function of a morpheme which is combined with other morphemes, e.g. *putoskop*. Interestingly, some of these morphemes, as e.g. *-logija* in (b1b) have, in their combination with other lexemes, taken on new meaning, new content, viz. ‘excessive amounts of’.

The different approaches and descriptions of lexicalisation supplement each other and sometimes overlap, The ‘elderly’-example and the *art*-example illustrate this overlap by being describable as an instance of *increase in autonomy*, i.e. a transfer from a syntactically dependent word category (adjectives) to an independent one (nouns) or as an instance of

conversion from one word category to another. The difference lies in the perspective. Whereas conversion focuses on the form, i.e. the word category, the proposed 'Cline d' focuses on the functional climb from syntactic dependency to independency.

A subcategory of lexical changes is called or *Relexicalisation*<sup>146</sup> or simply *Semantic change*. These terms refer to the ascription of new lexical meaning, i.e. new content to an existing content item. Language planners, proscribing to the so-called archaising purism may deliberately ascribe new meanings to old words instead of adopting foreign words (cf. "2.2.3 Purism" above) and subsequently promote the actualisation and implementation process by promoting the relexicalisation through different channels.

The two processes of lexicalisation and grammaticalisation are, as we can see, not mutually exclusive. On the contrary, a change will often subsume both a lexical (content-related) and a grammatical (functional) change and whether a process is viewed as lexical or grammatical depends primarily on the perspective. Is for instance the relatively new (and presumably contact-induced) *nana čaj* – 'mint tea' (T. Prčić 2006: 415) or *višnja čaj* – 'cherry tea' (Horvat and Štebih-Golub 2010: 13) to be viewed as an instance of lexicalisation, i.e. the formation of a compound, or should it rather be analysed as an example of an emerging new type of NPs in Croatian or Serbian and hence a grammatical innovation? In whichever way we label this innovation, we will have to recognise that an in Croatian and Serbian rarely seen syntactic or word-formational pattern is used more frequently than before. The 'more correct' equivalent to 'mint tea' and 'cherry tea' is *čaj od nane* and *čaj od višnje* ('tea out of mint/cherry')

Another field where both processes may be said to be in play is word-formation. Word-formation is traditionally dealt with in derivational morphology and hence a part grammar (e.g. (Bauer 1988: 125) though other scholars consider word-formation an aspect of lexicalisation<sup>147</sup>(e.g. Štekauer et al. 2012)

A word-formation subsumes a lexical change and this lexical change will in some instances also imply a grammatical change. The lexical change (in word-formational terms: conversion) of the numeral

---

<sup>146</sup> In a similar way, Andersen among others also calls the ascription of a new function to an existing function item: *regrammaticalisation* or *regrammation*

<sup>147</sup> lexicalisation, meaning the creation of words

*jedan* ‘one’ to the article *jedan* ‘a’ happens concurrently with the change in the grammatical system in the language from being a language without definiteness expressed through articles to being a language in which a reference to the grammatical category of definiteness is rarely omitted. *jedan* is at the same time grammaticalised as it goes from having the lexical content “1” to the grammatical function of marking indefiniteness.

The contact-induced innovations in Serbian and Croatian under investigation in this study concern content and function items that have been sufficiently actualised by ‘cohorts of speakers’<sup>148</sup> to *expand* into the written usage norm and have, except the emerging marking of indefiniteness been commented and advised against by both Croatian and Serbian normativists.

### 3.2.2 Types of innovation

Turning again to Andersen’s definition of language change we will now look at the different types of innovation that may occur in a language. Andersen lists four basic types and four contact types.

He applies a distinction between content and expression which I take to mean the semantic content and the form in which this content is expressed and defines four basic types (B) and four contact types (C)<sup>149</sup>:

(B1) *Neologism* which covers (i) new combinations of both content and expression, also called coinage, (ii) new ways of expressing existing content e.g. the introduction of new derivative affixes instead of using existing ones and (iii) new, derived realisations of existing expressions, e.g. clipping (as: *elektronapon* < *električki napon*). Andersen underlines the fact that a neologism will always be an alternative to an existing way of expressing the same content and will therefore in the initial stage be a variant of the existing one, e.g. *a gentle man* > *a gentleman*. Later, the novel expression, the innovative variant, may become the unmarked expression and curtail, i.e. alter the semantic content of the former expression.

(C1) *Borrowings*, which are kinds of bilingual *neologisms* which cover the introduction of signs from another language or the element by element translation of composite signs from another language, e.g. Serbian or

---

<sup>148</sup> cf. Andersen’s description of the process through which an innovation much go to become a language change, p. 116

<sup>149</sup> The following descriptions of Andersen’s types of innovation is partly paraphrased from Andersen (2008: 22–24) supplemented with Andersen (2001: 229–34)

Croatian *marketing* introduced from English 'marketing' and the Serbian and Croatian *djelokrug* introduced by translating the German compound 'Wirkungskreis' element by element ('Wirkung' > *djel*-, 'Kreis' > *krug*).

(B2) *Extension*, which covers the extension of existing linguistic means in new usage, e.g. the usage of the numeral *jedan* ('one') as an indefinite article ('a/an'). The extended expression will typically have less semantic depth than its source, i.e. it will typically be less concrete and/or referential. In this respect the article usage of *jedan*, (or the usage of English 'a/an') is again very illustrative as it still refers to something, of which there is only one item/person but its grammatical function of determining something as indefinite has more import than its numerical meaning

150.

(C2) *Intrusions* which cover *extension* from one language to another of new elements, or in my interpretation; induce new or increased usage of existing linguistic means. The article-like usage of the numeral *jedan* has been pointed out as a 'foreign influence', as well as has the increased creation of compounds such as *djelokrug* and later creations of compound-like constructions such as *nana čaj*.<sup>151</sup>

As can be seen *djelokrug* serves both as an example of *borrowings* and of *intrusions*. *djelokrug* is in itself a *borrowing*, but the introduction or increased usage of a noun as modifier of another noun is an example of *intrusion*. In other words, several of the four basic + four contact innovation types might well have been in play in the process leading to one single innovative expression.

(B3) *Adoption* simply covers the speaker's acceptance of newly encountered usage and his or her own actualization and implementation of this use. The actualization and implementation may, as we recall, lead to that the innovation becomes adopted by so large a number of the speakers that it becomes usage norm in the speech community, which eventually may lead to its acceptance and implementation into the standard language, cf. subsection 2.4.4 above. Whereas Andersen compares adoption

---

<sup>150</sup> In section 4.6 the semantic import of the English 'a/an' is defined according to Lyons (1998) and the article-like usage of *jedan* and *neki* are analysed and defined according to scholarly literature on the subject as well as on the basis of empirical data.

<sup>151</sup> *jedan* – (Maretić 1963 [1899]: 510), *djelokrug* – (Rožić 1904: 7), *nana čaj* – (T. Prčić 2006: 415)



to his four other types of innovation and focuses on the *How?* i.e. the functional aspect of adoption, sociolinguists are primarily concerned with the social motivation behind the adoption, i.e. *Who* (which language users) adopts what innovations from whom? Through *which channels* are innovations promoted and *why* the language users are compelled to adopt or indeed reject an innovation? Andersen's focus on the functional aspect does not mean that he regards adoption as a mechanical, involuntary, automatic process. On the contrary, he emphasises that adoption is, as opposed to the fourth type, reanalysis, a purposive process, meaning that speakers consciously adopt or reject innovations.

(C3) *Contact adoption* is the acceptance of alternative linguistic means from a contact dialect. Andersen does not concretise the term contact dialect, but I will interpret it as a similar idiom with which the innovative language is in contact. Again *djelokrug* may serve as an example as it was firstly introduced in the Croatian speech community as a borrowing and later adopted by the Serbian speech community.

In a sociolinguistic approach, i.e. when addressing the question of why a linguistic means is or is not adopted, it is interesting to see what is and what is not adopted into the usage norm because of the symbolic value inherent in the adoption or rejection of said means. I therefore propose to widen the definition from "the acceptance of alternative linguistic means from a contact dialect" to "the acceptance of alternative linguistic means from a contact language or dialect" thereby including adoption from any language or language variety with which speakers come into contact. This makes it possible to include the adoption of linguistic means other than the lexical ones covered by the Andersen's process of *Borrowings*.

(B4) *Reanalysis*, the fourth type concerns the way in which new speakers analyse and valorise existing content and expression. Reanalysis is what leads to grammatical change. Andersen clearly has first language acquirers (typically children learning their mother tongue) in mind (Andersen 2001: 232). However, contact-induced innovations often undergo a similar type of (re)analysis by speakers of the replica language, thereby narrowing or broadening the semantic scope (content) of the replicated expression.

(C4) *Bilingual reanalysis*. Andersen singles out the fourth contact type as one of *bilingual reanalysis* and explains that this term covers the composite competence of social bilingualism or bidialectism in which the speakers' valorisation of two equivalent elements from the two idioms determines the speakers' preference in usage. Even if we were to interpret *bilingual* as somebody who merely knows how to use the contact-induced linguistic matter or pattern in the model language, the rate of such bilinguals among Serbs and Croats is still so small that they, in my view, will not be able to constitute 'cohorts of speakers' and as such lead to the acceptance of an innovation. The bilinguals will often play the role of innovators but the implementation of an innovation will be dependent on the adoption of it by non-speakers<sup>152</sup> of the model language.

In an attempt to apply the already mentioned examples of innovations in Serbian and Croatian I am tempted to mention the noun-modification strategy in English which results in compounds composed  $N_{mod} + N_{head}$  which is supposedly the source of examples such as *nana čaj* where the older variant *čaj od nane* ( $N_{head} + PP_{mod} (P+N)$ ). The *bilingual reanalysis* would, in this instance, consist of the speakers' evaluation of the two variants, and result in a preference for one of the two variants.

In the adoption and reanalysis types of innovation, the innovations will exist alongside the existing way of expressing the same content. E.g. the innovated or incoming variant *nana čaj* exists alongside the older; outgoing variant *čaj od nane* as well as the functional article-like usage of *jedan* exists alongside the numerical content-usage. One way to ascertain whether this or other innovations are gaining ground in a speech community is to attest the frequency of the innovative, incoming variant (the innovation) as opposed to the older, outgoing variant. As the innovative variant gains in frequency, it loses its oddity or rather; its markedness, thus shifting from a stylistically or pragmatically marked position to an unmarked position.

Contact-induced linguistic innovations or in Andersen's words *externally motivated innovations* will via their occurrence inflict a change in the usage rules, i.e. in Coseriu's terms *the norm*, of said language. At first it will occur „in the most salient, monitored, marked environments, from which it may spread, as it loses its novelty, to less salient, unmarked environments.“ (Andersen 2001: 34).

---

<sup>152</sup> The „non-speakers“ of the Model-language need not be monolinguals, but are only defined here by not knowing the specific Model-language in question

Once again, it must be added that the spread of any innovation into the standard language may be blocked or inhibited by language planners. These language planners will in such instances through acquisition planning and codifying works on the standard language define and prescribe a standardised norm which is in conflict with an innovative usage among the speakers in the speech community covered by this standard language. The results of my analysis of the empirical data will show to what extent the innovations are present in the usage norms of Serbian and Croatian, regardless of the prescribed standardised norm.

The reasons for making the innovation are many, the motivation for the particular innovation likewise. Regarding the motivation, there has long been at least one division among scholars who focus on the linguistic innovations themselves. First we have (historical) linguists who concentrate on language internal motivations, i.e. the development of a new perfect tense due to grammaticalisation of full a verb (*jesmi*) to an auxiliary (*jesmi*) to an affix (-*m*) not their possible external motivation (cf. p. 120); and then we have (contact) linguists who focus upon the language changes which are externally motivated by contact with another language/other languages.

### **3.3 Contact-induced language change**

What is considered contact-induced language change?

All changes induced i.e. provoked or motivated through the contact between, on the one side speakers of the changed language and on the other the usage norm in the language of contact.

Thomason (2001), in *Language Contact – an Introduction* asks the following intriguing question: “[W]hy do some communities borrow foreign words along with foreign cultural items, while others create native words for cultural borrowings?” (Thomason 2001: 82). If one should attempt to answer this question, one would have to take into account a range of factors ranging from linguistic structural compatibility to attitudes towards the lending speech community. In other words, it is formally easier to incorporate elements from another language, if the phonetic, morphological, word-formational, syntactic, semantic and stylistic make-up of the other language is similar to that of the replica language, i.e. if they are typologically similar and/or genetically close to each other.

However, the replication of the salient features in the model language will have difficulties in being accepted and implemented by a large group of speakers of the replica language if the replicating speech com-

munity bears hostility towards the speech community of the model language or is generally suspicious of all foreign linguistic features as is the case in situations where reformist and/or xenophobic purism, is active.<sup>153</sup> I will, as does Thomason, regard this relationship as a conflicted relationship of linguistic versus social factors that all potentially facilitate or impede contact-induced language change.

### 3.3.1 Types of contact-induced innovations

In subsection 3.2.1 different approaches to lexical and grammatical change were outlined and in subsection 3.2.2 a general outline of four types of contact-induced innovations were introduced: *Borrowings*, *Intrusions*, *Contact adoptions* and *Bilingual reanalysis*

Among Andersen's four contact types of innovations only *Borrowings* and *Contact adoption* explicitly concern the replication of lexical matter (loanwords) as well as lexical pattern (calques) from a model language, while *Intrusions* concerns the extension of patterns and usage of existent matter in a way present in the language(s) of contact. The *Bilingual reanalysis* mostly has to do with the actualization, acceptance and implementation of the matter and patterns *borrowed*, *adopted* or *extended* from contact languages.

Scholars who have concentrated their work on contact-induced language changes provide us with a slightly different terminology. In the following section I will add insight and explanatory frameworks used by Thomason (2001; 2003), Heine and Kuteva (2005; 2006; 2010), Harris (2003) and Matras (2009; 2010).

Heine and Kuteva distinguish between *borrowing of form meaning units* and *grammatical replication* and provide the following overview of possible linguistic contact-induced phenomena, which they call "kinds of linguistic transfer":

- a) Form, that is, sounds or combinations of sounds
  - b) Meanings (including grammatical meanings or functions) or combinations of meanings
  - c) Form-meaning units or combinations of form-meaning units
  - d) Syntactic relations, that is, the order of meaningful elements
  - e) Any combination of (a) through (d)
- (Heine and Kuteva 2005: 2)

---

<sup>153</sup> cf. "2.2.3 Purism" above.

Harris, who seeks to explain syntactic change, deals with replication of a syntactic pattern into the replica language through the influence of a pattern found in a model language and labels this phenomenon *borrowing* (Harris 2003: 529, 532).

I apply the terminology used by Matras, who distinguishes between ***replication*** of linguistic matter and *pattern*.

Matras' definition of *replication*:

"the term *replication* capture[s] [...] that we are dealing not with issues of ownership or even direct imitation or duplication, but rather with the activity of employing an item, in context, in order to achieve a communicative goal." (Matras 2009: 146) Matras' understanding of the term *replication* is not limited to the replication of items, as the quote might indicate. Replication may also be "characterized as a change to an inherited structure of the 'replica' language, inspired by a structure in the 'model' language." (Matras 2009: 238).

Haspelmath, who refers to Matras' distinction between matter and pattern, tells us that matter replication<sup>154</sup> refers to "lexemes, or more precisely, lexeme stems, but sometimes just affixes and occasionally perhaps entire phrases" whereas pattern replication<sup>155</sup> "refers to the copying of syntactic, morphological or semantic patterns" (Haspelmath 2009: 38–39). Linguistic matter encompasses concrete, identifiable wordforms and morphemes, whereas linguistic patterns are the ways in which the units of speech are organised (Matras 2009: 148).

So, when an item (a wordform or a morpheme) is replicated, we are dealing with replication of linguistic *matter*, and when a structure is replicated, e.g. the compound structure in *djelokrug* and *nana čaj*, or the usage of an indefinite marker as in the use of *jedan* we are dealing with replication of a linguistic *pattern*.

In other words, pattern replication is replication of a structure whereas matter replication is replication of material. Heine and Kuteva's different kinds of linguistic transfer (quoted above) will also serve as a guideline in distinguishing between the different types of contact-induced innovations under investigation in this dissertation.

---

<sup>154</sup> which Haspelmath terms *material borrowing*

<sup>155</sup> which Haspelmath terms *structural borrowing*

Combining the distinction between linguistic matter and pattern with the distinction between lexical and grammatical change we end up with four general types of replication:

*Replication of lexical matter, Replication of lexical pattern,  
Replication of grammatical matter, Replication of grammatical pattern*

The advantage of Matras' categorisation is, first of all, that he uses the term replication and not borrowing, which illustrates the fact that an innovative feature, inspired by another language is not borrowed (or lent), neither is it an exact duplicate of the feature in the model language, but rather a replication, i.e. a new rendering of an item or feature. Secondly, the term replication indicates that there is an active agent ('cohorts of speakers') behind the innovation.

The difference between lexical matter and grammatical matter is consistent with the difference between content items (lexical matter) and function items (grammatical matter), introduced earlier (cf. subsection 3.2.1 above)

Furthermore, the difference between lexical pattern and grammatical pattern is consistent with the difference between content and function. E.g. when the content of a compound such as 'Wirkungskreis' is replicated but not its expression (the item itself), resulting in *djelokrug*, *djelokrug* is the replica of a lexical pattern, but when the compound creation is innovated in Croatian and Serbian, inspired by a structure in a model language (English), resulting in [N N]<sub>N</sub> compounds as *nana čaj*

As noticed earlier, in order for any linguistic innovation to manifest itself as a generalised change in the language of a whole speech community it has to be adopted and actualized repeatedly by scores of speakers. During this repeated actualisation of an innovation its usage will gradually shift from being rare to frequent whereby it gains an increased probability to enter the usage norm.

An initial stage in a contact-induced change (a contact-induced innovation) may be the increased usage of a hitherto infrequent pattern in a language. One example could be an increased usage of an otherwise or hitherto quite restricted use and formation of verbal adjectives/present participles such as *fascinirajuće* instead of the adjective *fascinantno*, probably due to the contact with adjectival use of the English present participle 'fascinating'. Another example could be the increased usage of the numeral *jedan* and other wordforms in an article-like function. The

usage and maybe even the semantic scope of these wordforms and the constructions of which they are a part are thus *extended*.

Heine and Kuteva term this contact-induced phenomenon *replication of grammatical use patterns* (Heine and Kuteva 2005: 40).

The properties of grammatical use patterns are:

- a. They are associated with some specific grammatical meaning
- b. They are recurrent pieces of linguistic discourse. Such pieces may consist of a clause, a phrase, or even a single form used in some specific context
- c. Their use is optional, that is, they need not be employed for the expression of that grammatical meaning.
- d. They are primary units figuring in the initial stage of grammatical replication

(Heine and Kuteva 2005: 41)

The grammatical use patterns are not new to the language in question. In other words *fascinirajuće* and *jedan* were used or were at least usable in the described manner before the contact occurred, but their increased usage may lead to full-fledged replication of grammatical pattern, i.e. a change in the linguistic system. Or; to rephrase the well-known words of Roman Jakobson (1959: 236): The increased usage of a particular grammatical use pattern may lead to that definiteness is something that *must* be conveyed and not only *may* be conveyed in Serbian and/or Croatian. This increased usage constitutes in itself a change, instigated by speakers' wish or need to use a grammatical pattern, available in another language. The contact-induced increased use of a grammatical use pattern, Heine and Kuteva regard as *a rise from a minor to a major use pattern*, and characterise the rise as follows:

*The rise of a major use pattern in contact-induced replication*

- a. An existing use pattern is used more frequently
- b. It is used in new contexts
- c. It may become associated with a new grammatical function

(Heine and Kuteva 2005: 45)

As with many of the observations, claims and conclusions regarding contact-induced language change, the claim that increased and broader use of some linguistic means may trigger a grammatical change is not only valid for externally motivated (viz. contact-induced) language change

but are also applicable when dealing with so-called internally motivated language change.

When focusing on contact-induced content items, i.e. lexemes, it is useful to add Haugen's distinction between substitution and importation. *Importation* concerns replicated lexical matter, whereas *substitution* is in fact avoidance of foreign lexical matter (Haugen 1950: 214, 230). Alternatives to replicated matter - substitutions come in different forms, which I will return to in detail in „4.2.2 Substitution of replicated matter“.

*Lexical change* happens in all languages and may or may not be induced by contact. Only very rarely are new words created ex nihilo, as e.g. *hobbit*. So, new items are either created out of matter existing within the language by combining existent lexical and/or grammatical matter or they are replicated from other languages. This does not mean that items created within the language may not consist entirely or partially of matter previously replicated from other languages or indeed be formed according to a lexical or grammatical pattern present in a language with which the language users (i.e. the innovators) have been or are in contact. The items may be invented and formed within the language and thus their formation need not be contact-induced, even though the bits and parts (the matter) has entered the language through contact. The latter formations, consisting of existent bits and parts, but formed according to a model in a language of contact, will be considered contact-induced changes of the lexical pattern or word-formational changes.

Pattern replication affects the linguistic system, the structure of a language, whereas matter replication does not. One may also view these different effects of the types of replication as an example of ‘adaptation’ vs. ‘integration’. Replicated matter is *adapted* into the replica language, phonetically, morphologically and does not corrupt, i.e. change the system. Replicated pattern is *integrated* with the system, which means that it enters the existing system thereby changing it. Furthermore, when patterns are replicated, the existing system of the RL may change or an increased use of a hitherto less used pattern may be motivated. So, a new pattern or *a rise from a minor to a major use pattern* may very well be motivated by language contact and also via a preceding replication of matter from another language, viz. the model language (ML).

An increase in a RL of matter replicated from a ML in which a particular pattern is in use may therefore be seen as a likely trigger of an increase in the usage of a specific pattern in the RL. A consequence of the close connection between replication of matter and replication of pattern



is that the division into pattern and matter is not a division which is immediately applicable in form of a categorisation of the empirical data. Therefore, some of the investigated innovations as examples of as well matter as pattern replication.

Some scholars claim, contrary to my view, that replication of matter may not result in pattern replication. B. Brborić (2001) claims that one may use different lexemes (domestic in ( $S_{dom}$ ) or replicated/international in ( $S_{int}$ )) but the syntax will stay the same. B. Brborić illustrates this with the following two sentences, which I have translated into the English in ( $S_E$ ).

$S_{dom}$ ) *Polet i oduševljenje našeg naroda bili su mu osnovna snaga u neprekidnoj borbi za opstanak i oslobođenje*

$S_{int}$ ) *Elan i entuzijazam naše nacije bili su joj bazična energija u permanentnoj kampanji za egzistenciju i emancipaciju*

(B. Brborić 2001: 198)

$S_E$ ) 'The zeal and enthusiasm of our nation were its basic energy in [its] permanent struggle for [its continued] existence and emancipation.'

B. Brborić thus seeks to show that no matter how much you internationalise the lexical matter, the grammatical (syntactic) pattern stays the same. B. Brborić may very well be right in his assumption that replicated lexemes, which have been adapted, phonologically and morphologically may not change the topology of the R-language, but I will propose that replicated lexical matter may serve as a conductor of 'foreign' patterns; phonetic, phonological, word-formational, morphological, syntactic and semantic patterns which may enter Serbian or Croatian, i.e. be replicated via the replicated matter.

*Pattern replication* undoubtedly evokes a more fundamental change of a linguistic system, than the mere replication of lexical matter. In other words a change of the lexis is not a systemic change, whereas a change of structure is.

Schematically, some of the different parameters introduced so far are applied in Figure 3.1 below, in the description of the innovative variant *nana čaj* and the older variant *čaj od nane*.

Figure 3.1 - From nominal phrase to nominal compound

Innovation	Sociolinguistic variance		Recent linguistic change				Contact-induced
	Codified	Puristic evaluation	Innovation type	Use pattern	Variance	Marked	
$[N+ATTR]_{NP} > [N_{UNDECL}N]_N$							
<b>čaj od nane</b>	Yes	Correct		Major	Older variant	No	No
<b>nana čaj</b>	No	Incorrect	Borrowings/ Intrusion	Minor	Innovative variant	Yes	Yes

Depending on who (or whose usage) defines the norm, *nana čaj* and similar constructions could be on their way up from Silić's Coseriu-inspired concrete level of *speech* via the least abstract level *usage* and on towards the more abstract level of *norm*. But, because this construction is yet to be accepted by codifiers it has not reached the third level of abstraction: *codification*. However, the lack of codification, does not, in my view, mean that the  $[N_{UNDECL}N]_N$ -construction has not reached the fourth level of abstraction, the *system*.

In other words, I do not consider the codification an abstraction, but rather a deliberate and very concrete attempt at controlling the norm, thus creating a prescribed, standardised norm which is not restricted by the language system, but by the codifiers' understanding of the system. To put it differently, the system may change as a consequence of the change of usage norm, independently of the codification.

As indicated in Table 1, above, at least two of Andersen's types of innovation are in play: *Borrowings* and *Intrusion*. If *nana čaj* is interpreted as an „element by element translation of composite signs from another language“, it neatly falls under the innovation type *Borrowings*. If we focus on the pattern within the construction  $[N_{UNDECL}N]_N$  the innovation should be labelled *Intrusion* as the word-formational/syntactical pattern  $[N_{UNDECL}N]_N$  has been replicated into the RL, thus introducing a new pattern in the usage norm, allowed by the system of the RL which, through its occurrence, competes with existing syntactic ways of conveying the same semantic relation between the two nouns, i.e.  $[N+ATTR]_{NP}$  - *čaj od nane*.

Let us now apply the same template to illustrate the increased usage of *jedan* as an article-like marker of indefiniteness.

Figure 3.2 - From numeral to indefinite article

Innovation	Sociolinguistic variance		Recent linguistic change				Contact-induced
Quantifier > Indefiniteness	Codified	Puristic evaluation	Innovation type	Use pattern	Variance	Marked	
numerical: <i>jedan</i>	Yes	Correct		Major	Older variant	No	No
indef. marker: <i>jedan</i>	-	Incorrect	Extension/ Intrusion	Major	Innovative variant	No	(Yes)

A numeral, like *jedan* – ‘one’ refers to a particular amount and thus it has a lexical content as in (A), i.e., it is a content item. On the other hand we have the article-like *jedan* as in (B) which is a function item, more specifically a grammatical word.

1) *Nažalost, u Dubrovniku sam samo jedan dan* – ‘Unfortunately, in Dubrovnik I am just one day’ (HNK 2.5)

2) *Nazvao je prijatelje, novac su trošili na alkohol i drogu, a uhićen je 26 sati nakon pljačke kod jednog prijatelja.* – ‘He called his friends, the money they spent on alcohol and drugs, and he was arrested 26 hours after the robbery at a friend’s.’ (HNK 2.5)

The article-like usage of *jedan* – ‘a/an’ in (2) is thus an example of grammaticalisation according to the cline of grammaticality (Cline *a*, p. 120), where the function item usage (*jedan* – ‘a/an’) of a content item (*jedan* – ‘one’) exemplifies the climb from content toward function as it is used as a grammatical word:

Cline *a*

content	function item			
	grammatical	> clitic	>	inflectional

Today, the article-like usage of *jedan* is codified in normative works. This usage is recognised as that of an indefinite pronoun (Barić et al. 1995: 208) which is used to ‘express uniqueness combined with indefiniteness’ (Piper and Klajn 2013: 94, 114) or as a ‘determiner which may be used as an indefinite article’ (Mrazović 2009: 288). Earlier, it was deemed incorrect to use *jedan* as an indefinite article and it was also recognised as a contact-induced phenomenon:

'many of today's writers corrupt the language by the completely unnecessary use of the numeral *one* in accordance with the German article *ein*, French *une* and Italian *uno*'  
(Maretić 1963[1899]: 510; 1924: 37–38)<sup>156</sup>

But in accordance with contemporary grammars of both Croatian and Serbian, *jedan* as numeral and *jedan* as an indefinite article-like item are identically labelled, i.e. both uses are codified, both uses constitute a *major usage pattern* and are thus both *unmarked* vis-à-vis each other. By writing "(Yes)" in the column "Contact-induced", I wish to indicate that, despite Maretić's recognition of this phenomenon as a result of contact with German, French and Italian and despite some similar claims by Heine and Kuteva (2010) (which we will return to in section 4.6, below) there are Croatian linguists who claim that indefinite article-usage is a particularly Croatian phenomenon and not a contact-induced feature. According to the web-based encyclopaedia "Wikinfo" both indefinite and definite articles are in use in the dialects Kajkavian and Čakavian but have been suppressed by Maretić and other Vukovites during the union with Serbian.<sup>157</sup> In a very colourful way the anonymous authors of the encyclopaedic article tell us that: 'The use of an article is therefore a feature which is in the 'spirit of original Croatian' as opposed to that of other Slavs' and it would have been a part of Standard Croatian of today 'had it not been for the forced Serbo-Croat Yugo-union and 'rape' of the language the independent and natural Croatian standard would surely have officially regulated the use of the Croatian articles by the mid-20<sup>th</sup> century'<sup>158</sup>. While the colourful language, the anonymous authors and the fact that the article also claims that it is characteristic of a civilized people to have articles (while at the same time stressing that Serbian does not hold articles) strips the claims of credibility, it is also a fine example of how differences between Croatian and Serbian are sought out, maybe

---

<sup>156</sup> Original: „mnogi današnji pisci kvare jezik upotrebljavajući broj *jedan* bez ikakve potrebe prema njemačkom artikulu *ein*, franc. *un*, tal. *uno*,”

<sup>157</sup> See also p. 92ff for a discussion among members of „Council for Standard Croatian Language Norm“ of the claimed possible hybrid nature of the base dialect of Croatian.

<sup>158</sup> Original: "Uporaba člana pred imenicom je stoga imanentna osobina izvornohrvatskog "jezičnog duha" nasuprot inih Slavena" [...] „Da nije bilo toga prisilnog srbohrvatskog jugo-jedinstva i jezičnog 'silovanja', onda bi samostalni i prirodni hrvatski standard negdje do sredine 20. st. jamačno već službeno ozakonio izvornu uporabu naših izvornih imenskih članova." (Wikinfo 2016)

even thought up, and the homogeneity and sameness of Serbian and Croatian renounced.

Radovanović (2000: 28), who has observed an increased article-like usage of *jedan* in Serbian, sees it as a contact-induced phenomenon but not induced by contact with what he calls European languages. Instead, he sees it as an example of a balkanism, i.e. something evolving within the Balkan language league. Therefore the innovation type may be either *Extension* of an existing linguistic means in new usage or an *Intrusion* which covers extension from one language to another of new elements' (cf. p. 129).

### 3.3.2 Hierarchies or What is most replicable?

In the research on replication of matter (lexical and grammatical) several hierarchies or clines have been proposed. Field (2002) and Matras (2009) compare the different approaches, including the clines of grammaticality and lexicality introduced earlier (cf. p. 120).

It has been shown and claimed that some word classes are more easily or more frequently replicated than others. (Field 2002: 34–36), (Matras 2009: 153–65). Nouns are, due to their referential stability most readily replicated followed by verbs, adjectives and adverbs. (Matras 2009: 154). The study *English in Europe* (Görlach 2004: 7) claims that over 80 per cent of all replicated matter (from English into 16 different RLs) consists of nouns. Furthermore, it has been established that content items (nouns, verbs, adjectives and adverbs) are more readily replicated than function items, such as determiners, agglutinating (derivational) affixes and fusional (inflectional) affixes, (Field 2002: 38).

Thomason (2001) puts forward a *Borrowing scale* to account for what is mostly replicated depending on the intensity of the contact between languages (more accurately contact among speakers of different languages). On the second step of her four-step scale, she describes a situation where the contact is realised through a small group of reasonably fluent bilingual speakers of the RL who are a minority among the RL speakers. In this situation it is likely that content items and even function items are replicated, but there will only be a few instances of replication of grammatical patterns which could be new functions and new functional restrictions for existing syntactic structures, or increased usage of previously rare word orders. Even phonological features may be replicated but only as parts of replicated content items (Thomason 2001: 70). The present contact-induced innovations indicate that Serbian and Croatian are currently at this second step, at least when English is the ML.

Besides the instances of innovations under investigation in this study which fall into the group of replicated content items (*marketing* (lexical matter), *internet stranice* (lexical matter + pattern), replicated word-formational patterns (*nana čaj*) and new functions for existing syntactic constructions (*jedan prijatelj* 'a friend' < 'one friend'), phonological features such as previously 'impossible' vowel and consonant clusters have also entered Serbian and Croatian via replicated lexical matter. No two vowels were adjacent to each other if they were not part of different morphemes<sup>159</sup>, as e.g. *neuk* (ne-uk) ('unlearned'), but with replicated matter it became possible: *neatlantski* ('Neo-atlantic'), *aerodrom* ('airport') (Petrović 1996: 100–101). Consonant clusters in word-final position were not seen in Slavic Serbian vocabulary unless the clusters were *st*, *št*, *zd*, *žd*. Many other clusters are seen in replicated matter, thus being an innovation in the structure of consonant clusters in Serbian and Croatian, e.g. *koncept*, *pakt*, *lift* ('concept, pact, lift') (1996: 105).

On the third step on Thomason's *Borrowing scale* the users of the RL replicate more grammatical matter, as e.g. derivational affixes and pronouns, even inflectional affixes. Phones not present in the model language are not realised anymore, replicated phones are used in non-replicated matter. Word order and syllable structure is changing etc. (Thomason 2001: 70). Serbian and Croatian have not reached this stage of replication except perhaps regarding the replication of derivational affixes which in some cases were replicated not from English, originally, but rather from Latin (via German and Old French): *-ira-* as in *definirati* ('to define') and from Greek: *-isa-* as in *definisati* ('to define').

### 3.3.3 The assumption of contact

When pattern innovation is concerned, it is virtually impossible to prove that the specific pattern innovation is in fact a pattern *replication*. I assume that *internet korisnik* is a replication of 'internet user', because the [N<sub>UNDECL</sub>N]<sub>N</sub>-construction has been deemed not in accordance with word-formational regularities within Serbian and Croatian. And I assume that the model language is English because I know that English is the current dominant foreign language in Serbia and Croatia as in a large part of the world. I cannot prove it but I, as well as many other students of contact-induced change and contact linguistics, do assume that the trigger effect

---

<sup>159</sup> Except when the result of the word-final or syllable-final alternation l>o as in *pepeo* ('ashes'), *seoba* ('migration')

for the changes which I investigate is contact. But, *internet korisnik* and for that matter all other instances of [N<sub>UNDECL</sub>N]<sub>N</sub>-constructions could also be interpreted as a language internal change, i.e. an innovation triggered by something else than contact. It is, unquestionably, scientifically problematic to assume and not prove that a given linguistic innovation is contact-induced. Thomason (2001), having underlined the fact that a linguistic innovation stays a linguistic innovation no matter what triggered it, points to some necessary requirements for assuming that an innovative pattern or an increased use of a pattern is indeed motivated by contact to another language. First of all, the presence of immediately identifiable contact-induced innovations in the replica language, i.e. matter replication, is necessary. In other words: It is safer to assume pattern replication from a given model language if matter is replicated from the same model language. Secondly, having identified the model language, the possibly replicated patterns need to be identified in both languages. Thirdly, we must prove that the proposed replicated patterns did not exist (or were in minor use) before the proposed replication took place and that the patterns did/do exist in the model language. (Thomason 2001: 93–94). This still does not prove that a given change in a language was contact-induced but it heightens the probability of it. In this study Thomason's requirements for assuming contact as an inspirational factor for linguistic innovations in the proposed replication of the word-formational pattern [N<sub>UNDECL</sub>N]<sub>N</sub> and in the proposed emergence of an article-like marker of indefiniteness are met in as much as there are many instances of matter replicated from English into Serbian and Croatian. The patterns in question are definitely present in English (e.g. *internet user* [N<sub>UNDECL</sub>N]<sub>N</sub>, *a friend* [ART<sub>INDEF</sub> N]<sub>NP</sub>). Proof that the patterns were non-existent in Serbian and Croatian before the contact does not exist, but the increased use of these patterns does coincide with the increased contact with English. The reactions from the linguistic authorities in the two speech communities, i.e. to which innovations they react (negatively) and which language they presume to be the inspirational factor is ample proof of not only the existence of these innovations but also of their implementation in the usage norms of the languages.

In my investigation of linguistic innovations in Croatian and Serbian, I will be analysing instances of lexical matter replication (of nouns and nominal compounds), bound lexical and grammatical matter replication (semiwords and affixes), word-formational pattern replication (nominal compounding), and finally an instance of *a rise from a minor to*

a major grammatical use pattern, which could lead towards grammatical pattern replication (of indefiniteness marking). I assume that the word-formational pattern replication (in nominal compounding) is induced by the existence of replicated lexical matter (compounds). I assume this because I am convinced that the replicated linguistic matter also (covertly) provides a loop-hole for linguistic pattern through which it may seep into the RL, i.e. be extended from the replicated matter into the inherited matter, thus triggering a change in how matter is organised in the RL. To signal the fact that the pattern is not replicated directly from a pattern in the ML but rather replicated or extended from already replicated matter I propose to term it *relayed pattern replication*<sup>160</sup>.

In the following chapter I will proceed to describe, problematize and define the concrete phenomena under investigation, which are: replicated *nouns* and their suggested substitutes, replicated *compound structures*, replicated *semiwords* and *affixes* and replicated *indefinite marking*. Subsequently, I will present methods used and results gained from my research into the empirical data found in the Croatian and Serbian 100 m+ national electronic text corpora and relate my findings to existing research as well as to the preceding chapters.

---

<sup>160</sup> 'relay' in this term is inspired by the term *Relay interpretation* which means interpreting between two language via a third.



#### 4. Investigated Linguistic Innovations

All four general types of replication: replication of lexical matter, replication of grammatical matter, replication of lexical pattern and replication grammatical pattern are relevant in the study of the selected linguistic innovations in Serbian and Croatian. The replicated matter (lexical or grammatical) will undergo certain formal modifications when replicated. This is not the case with replicated patterns as they, per definition, have no form.

Henrik Gottlieb in his article “Linguistic Influence” (2006) provides an overview of standard arguments for and against replicating lexical matter:

<u>Pro</u>	<u>Con</u>
Facilitates learning of the model language	Impedes reading of national classics
Shortens distance between languages and cultures	Increases distance between generations and social groups
Provides expressive enrichment	Leads to linguistic impoverishment
Makes translation simpler	Kills the fascination with foreign languages and cultures
Fights chauvinism and provincialism	Paves the way for foreign cultural dominance
(Gottlieb 2006: 197)	

As differences in the lexicon between Croatian and Serbian have been the focus of much research since the break-up of Yugoslavia, there has also been conducted research into how the standard languages and the language usage is affected by contact-induced lexical innovations, i.e. as to the acceptance and implementation of replicated lexical matter and pattern.

##### 4.1 Recent research in lexical divergence between Serbian and Croatian concerning replicated matter

Even though there has been much dispute as to whether Croatian and Serbian should indeed be considered two different languages it is commonly accepted that there is and always has been a lexical difference between the two. Most of the work on comparing the two has been motivated by a desire to underline the differences or rather to underline what is Croatian and what is not Croatian and to guide or teach the Croats to

use proper Croat words instead of Serbo-Croatian (i.e. Serbian) words). This work has been done chiefly through the writing of new dictionaries, orthographic manuals and language guides in Croatia.

Since Serbo-Croatian ceased to exist as a *lingonym* in Croatia and Serbia, scholarly work with an equal focus on both variants and attempts at comparing the two have been scarce if we do not count the quasi-academic Croatian publications of vocabulary lists and so-called differential dictionaries „...which explicitly focus on that which is Croatian and that which is not Croatian (i.e. which is considered to be Serbian).“ (Alexander 2006: 402). The differences between what is considered and to some extent codified standard Croatian and standard Serbian are constantly underlined, but there is very little research based on the actual language *usage*<sup>161</sup> with a focus on the differences. As a teacher of Bosnian/Croatian/Serbian, I can testify, that serious research into this particular subject is sorely missed, not so much as regards the differences between the normative standards but more so with a view to difference in usage norm in the five functional styles. There is, understandably, even less research on the possibly different contact-induced linguistic innovations in Serbian and Croatia of today. It is, though, as mentioned above, often repeated that Serbian is more open to foreign influence than Croatian, that is, the Croatian language develops under the influence of puristic attitudes whereas the Serbian language does not.

Having compared the research done in Croatia and Serbia on contact-induced linguistic innovations, I detect a striking similarity in the described innovations, which was what led me to assume that a comparison of empirical data from Serbia and Croatia would show that the same types of innovations are present in both languages, but at the same time I expect a difference in the acceptance and implementation of these innovations.

A project, led by Branko Tošović at the *Institut für Slawistik der Karl-Franzens-Universität* in Graz, Austria had as its goal to investigate the differences between Bosnian, Croatian and Serbian. It resulted, among other things, in four anthologies, published in 2008 and 2009, of which the second (Tošović 2009) focused on lexis, word-formation and phraseology with a special emphasis on replication (Tošović 2009: 2: 13). In the anthologies, linguists from Serbia, Bosnia-Herzegovina and Croatia

---

<sup>161</sup> *Usage* is, as we may recall, one of the three elements, which together with *codification* and *norm* that, according to Silić, constitutes the standard language (cf. p.67)

as well as a handful of other countries contribute with various articles addressing either the differences between two or all three of the languages or addressing specific developments in one of the three languages.

In this anthology Rada Stijović from Belgrade tells us that the lexical norm of standard Serbian has not changed since the break-up of Yugoslavia and that Serbian lexicographers still principally regard Serbian and Croatian as one language and therefore continue to include lexemes which during the Serbo-Croatian period were regarded as part of the western (i.e. Croatian) variant of Serbo-Croatian. (Stijović 2009: 219–20). Stijović's claim corresponds very well to my impression of the Serb view of Serbian in which Serbian is perceived as the natural successor to Serbo-Croatian in the same way as the Serbs view themselves as the natural caretakers of the Yugoslav spirit.

Milica Vasyľeva from L'viv, which in her contribution focuses on tendencies on the lexical level of Serbian, Croatian and Bosnian, claims on the basis of an investigation of contemporary dictionaries, grammars and language manuals as well as magazines and newspapers from Serbia, Croatia and Bosnia-Herzegovina that the standardised lexis of all three languages has undergone a true revision due to extra-linguistic factors since the separate standardised norms were established, that is, since they gained the status of separate standard languages. She concludes, as have many others, that the Croatian standard is more closed to foreign lexis as well as what is considered Serbian lexis, whereas Serbian is more open to as well foreign lexis as Croatian lexis. Furthermore, she provides lists of lexical innovations in Croatian which have been triggered by the need to substitute foreign lexis with domestic. (Vasyľeva 2009: 165–75)

Lelija Soćanac from Zagreb tells us, on the basis of an investigation of a corpus of text from the printed media and web pages<sup>162</sup>, that in the observed language usage, despite the glottopolitical, puristic official objections and campaigns against Anglicisms in Croatia, there is no discernible difference between Serbian and Croatian as to the quantity and quality of replicated lexical matter from English. The differences lie in the adaptation of the replicated matter, not in whether the lexical matter is replicated or not (Soćanac 2009: 250).

---

<sup>162</sup> The author, unfortunately, does not date these sources, but I believe them to have been published after year 2000

Branka Drljača Margić from Rijeka (Croatia) compares the use of English vocabulary in advertising in Serbia and Croatia and her conclusion is, similarly to that of Sočanac, that despite the Croatian purism, the use of English in advertising is equally present in the investigated advertisements in five women's magazines, published in both countries in September and October 2007<sup>163</sup>. (Drljača Margić 2009: 266, 276)

Milica Mihaljević and Barbara Štebih-Golub (Zagreb) compare Croatian and Serbian computer terminology by consulting written media, specialist dictionaries, text books in ICT and web based fora for IT-specialists, and conclude that a) there are many Anglicisms in both languages, b) Anglicisms are orthographically adapted in different ways, but morphologically in the same way, c) in Croatian in the scientific functional style more neologisms are used instead of loanwords than is the case in Serbian, d) when semantically replicating a foreign term, either through calquing or loan rendering, there is in both languages a tendency to translate one word terms with multiple words, e.g. *softver* 'software' > Ser. *programska podrška* ('programme support') / Cro. *računalna podrška* ('computer support') (Mihaljević and Štebih-Golub 2009: 141)

Milorad Dešić (Belgrade) concludes after having compared Croatian daily newspapers from Croatia with Serbian from Serbia, Montenegro and Republika Srpska, one issue from each country, all from 2008, that a) foreign lexis is more used in Serbian than in Croatian, b) there is a larger number of neologisms and revived archaisms in Croatian, c) the usage of lexis in Serbia, Montenegro and Republika Srpska differ very little from each other. (Dešić 2009: 117)

Ljudmila Popović (Belgrade) has investigated electronic discourse, i.e. discourse via e-mail, SMS, chat rooms and similar fora among Serbs and Croats and does not observe any dramatic difference between Serbian and Croatian use and adaptation of Anglicisms. In both groups of speakers there is a) a tendency to transcribe the English words and/or use transcribed versions of Anglicisms, e.g. *apgrejd* < 'upgrade', *mejl* < 'mail', *onlajn* < 'online, b) a tendency to apply domestic verbalising suffixes to English verbals stems, e.g. Ser. *restartovati*/ Cro. *restartati* < 'to restart', Ser. *aploudovati*/ Cro. *uploadati* < 'to upload', Ser. *anzipovati*/ Cro. *unzipovati* < 'to unzip' in which you see the Serbian preferred verbal

---

<sup>163</sup> 60,9% of the advertisements in the Croatian magazines and 61% in the Serbian magazines out of a total of 497 advertisements contained English

formant *-ovati* (*restart-ovati*, *aploud-ovati*) and the Croatian *-ati* (*restart-ati*, *upload-ati*), but also c) a tendency to adapt the spelling more to pronunciation in Serbian electronic discourse (*anzipovati*, *aploudovati*) than in Croatian (*unzipovati*, *uploadati*) (Lj. Popović 2009: 199)

Arno Wonisch (Graz) has, on the basis of the empirical data from the Croatian National Corpus, Corpus of Contemporary Serbian<sup>164</sup> and the smaller Gralis-corpus which also includes Bosnian, compared specific parts of the sports terminology used in the media in Bosnia, Croatia and Serbia. He concludes, on the basis of a limited sample of words collected from articles on three different sports: football, tennis and skiing, that there are several differences, mostly between Croatian on the one side and Serbian and Bosnian on the other. Most of the listed lexemes are English loanwords (replicated lexical matter) or neologisms (substitutions). In Serbian and Bosnian the replicated matter is used to a larger extent than in Croatian, e.g. Serbian and Bosnian: *ofsajd* < 'offside', Croatian: *zaleđe* - offside ('behind back'). Wonisch also speculates, on the basis of his investigation, that terminology in a sport which is known by everybody (as is football) the likelihood of different terminology, i.e. elaboration of own terminology, is larger than in narrower and more elitist sports such as tennis and skiing. (Wonisch 2009: 205–11).

In 2004, Croatian linguist Anita Skelin Horvat published an article on lexical replication into Croatian in two periods<sup>165</sup>, in which she, on the basis of an investigation of daily newspapers in the period 1970-1975 and 2000-2004 respectively, concludes that internationalisms (lexis from Latin and Classical Greek) are much less present in the 2000-2004 than in the 1970-1975-corpus: 65% (1302 out of 1988) in 1970-75 as opposed to only 16% (124 out of 759) in 2000-2004. For Anglicisms the situation is opposite: a high percentage of Anglicisms – 74% - in 2000-2004 as opposed to a low percentage in 1970-75 – 11%. (Skelin Horvat 2004: 101). The author deduces that the low percentage in internationalisms in the 2000-2004 is due to the expressed and partly successful purism in 1990s, which was targeted at serbisms and international lexis (also) used by Serbs. However, the puristic efforts to cleanse the Croatian language of Anglicisms showed to be far less successful. (Ibid: 99, 102)

The prominent Croatian linguist, Ivo Pranjković, lists in his article "Normative und paranormative Neuerungen in der kroatischen Sprache"

---

<sup>164</sup> The same corpora used in this study.

<sup>165</sup> Original title of article: *Posuđivanje u hrvatski u dvama razdobljima*

(2000) what he considers numerous innovations in Croatian which occurred during the 1990s. He ascribes these innovations to a number of socio-political reasons as e.g. the shift in relations to the Serbian standard language, the regime change and the subsequent establishing of the state institutions in Croatia, including the army, which happened following the Croatian declaration of independence from Yugoslavia, the radicalisation of language politics among scholars and the budding interest in language issues among laymen (Pranjko<sup>vić</sup> 2000: 66). Pranjko<sup>vić</sup> states that a considerable number of internationalisms have almost disappeared from the administrative publicistic styles of standard Croatian and lists 54 internationalisms and other foreign lexemes and their Croatian substitution (Ibid: 71). So, the results of Skelin-Horvat (2004)'s statistical empirical study (cited above) supports Pranjko<sup>vić</sup>'s claims.

In Serbia, a conference was held as early as 1995 on „Foreign words and phrases in the Serbian language, with regard to the same problem in the languages of the national minorities“ An anthology of papers held at this conference with the title „On lexical borrowing“ was published in 1996 (Planko<sup>š</sup> 1996). The Serbian attitude to foreign lexical matter is summed up nicely in Branislav Brbori<sup>ć</sup>'s contribution: 'The relation towards loan-words: With moderation but without aversion'<sup>166</sup> That is, replicated matter should be accepted but not without due consideration as to whether the items have domestic equivalents and whether these domestic equivalents (including calques) should have precedence over the replicated matter or not. Brbori<sup>ć</sup> advocates the thought that replicated matter which is less transparent to the average speaker should have domestic doublets (calques) (B. Brbori<sup>ć</sup> 1996: 46). Bugarski (1996: 24) also underlines the necessity of replicated matter, but that it should be used moderately and without prejudicial judgements along the lines of good/bad, pro/con.

Egon Fekete, on the other hand, advocates a very open attitude to replicated lexical matter and, in his contribution, welcomes replicated matter, claiming that resisting it is irrational and a sign of resistance to all things foreign, an attitude which is harmful to the language, and which he exemplifies with examples of Croatian substitutions of replicated lexical matter. On the contrary, replicated matter, is a source of enrichment of the language as it forwards its main function: communication (Fekete

---

<sup>166</sup> Original: "Odnos prema tuđicama: sa merom ali bez averzije"

1996b: 54). As noted above, in the section on purism (“2.2.3 Purism”, 76ff), the very open attitude towards replicated matter may even be interpreted as a deliberate and therefore symbolic attempt among Serbs to distance themselves from what they consider a Croat tendency (Klajn 2008: 157). In other words by being *antipuristic* the Serb language users signal their ‘anticroatism’, or put more mildly, their desire not to speak as the Croats do.

Željko Čupić has investigated cultural columns in 43 issues of the Serbian daily newspaper *Politika* in 1995 and concludes that there is a relatively large number of, what he calls recently replicated lexical matter, primarily nouns but also nominal compounds. Čupić concludes that the cultural columns are full of replicated lexical matter of older and more recent date but refrains from judging whether this is a positive or negative development. He adds, though, that the use of the replicated lexical matter is in most cases ‘moderate, pertinent and reasonable’.<sup>167</sup>(Čupić 1996: 317). So, Čupić like his fellow Serb colleagues and co-contributors to „On lexical borrowings“ does not object to replicated lexical matter as long as it is used with moderation and in a sensible way.

Among the contributors to „On lexical borrowings“, although they are in the minority, we also find a more antagonistic attitude towards replicated lexical matter. Ljubica Prčić, who analyses the meaning and connotations of replicated lexical matter which is used to characterize human characteristics, ends her article by announcing that ‘out of the torrent of foreign words we (the Serbs) should only use the ones of which we do not have domestic equivalents and that the rest feel like *intruders* and are a product of faddishness, snobbism, quasi-learnedness and lead to impoverishment of our (the Serbian) language.’ (L. Prčić 1996: 151)<sup>168</sup>

As a result of a project named *Modern Changes in the Slavic Languages*<sup>169</sup> initiated by Stanisław Gajda at Opole University in Poland in 1992, 14 volumes, each about one Slavic language, written in that Slavic language was published between 1996 and 2004. The title of the series is: *Najnowsze Dzieje Języków Słowiańskich* (‘The Recent History of Slavic Languages’)

In the Serbian volume *Srpski na kraju veka* (‘Serbian at the end of the century’) edited by Milorad Radovanović (Radovanović 1996), eight

---

<sup>167</sup> Original: “... umerenom, umesnom, razložnom upotrebom strane leksike.”

<sup>169</sup> Polish original: *Współczesne przemiany języków słowiańskich*.

distinguished linguists from Serbia describe Serbian and its sociolinguistic and linguistic development chiefly in the period 1945-1995. In the chapter on lexis „Leksika“ by Ivan Klajn, the author explains how replicated lexical matter is the Serbian vocabulary's primary source of enrichment as 63 pct. of the entries in the author's own dictionary of new words, Klajn (1992) *Rečnik novih reči* consist of loanwords (Klajn 1996: 46). A large majority of these loanwords have been replicated from English (Ibid: 45). Klajn does not express any negative or positive attitude towards replicated lexical matter. He does, however, claim that the reason for the overwhelming number of loanwords is not the dominance or 'torrent' of foreign words, but rather the inadequate creative power within the domestic vocabulary (Ibid: 46)<sup>170</sup>.

In the Croatian counterpart to 'Serbian at the end of the century' *Hrvatski jezik* ('The Croatian Language'), edited by Mijo Lončarić (Lončarić 1998), 16 Croatian contributors describe Croatian and its sociolinguistic and linguistic development in the period 1945-1995 but a lot of space is also reserved for explaining and emphasising the importance of the history of the Croatian language before 1945.

In the chapter on lexis, „Leksik“ by Marko Samardžija, the emphasis is on the forced 'serbification' of the Croatian version of Serbo-Croatian, the suppression of Croat puristic tendencies and the general stigmatisation of Croatian vocabulary. At the end of the chapter, Samardžija explains how the Croatian linguists since the spring of 1990 have been addressing three problems: 1) the problem of the loanwords, particularly the English ones, i.e. how to adapt them and/or how to substitute them with domestic lexical matter, 2) Serbisms in Croatian; i.e. how to identify them and abolish them and finally 3) explain to Croats that not all Croatian lexis, which had been marginalised since the Vukovites won the glottopolitical battle within Croatia in 1889 (cf. p. 16), should be revived because some of it has simply become obsolete (Samardžija 1998: 148–49). I will return to how lexis is de- and promoted by Croatian linguistic authorities in section 4.3, below.

On basis of the preceding preview of recent research in lexical development in Croatian and Serbian we may conclude that even though there is a focus on the differences in the codified standard languages, the research into actual developments, i.e. linguistic innovations present in the

---

<sup>170</sup> Original: "... nije u pitanju "naježda" ili "poplava stranih reči"[...] nego nedovoljne stvaralačke moći domaćega vokabulara"



language *usage*, for the most part suggest that the languages, Serbian and Croatian, in spite of clear differences between Croatian and Serbian normativist attitudes, develop in a similar way when it comes to changes and innovations induced by contact with English, the prestigious, global language of today. We may also conclude that it is difficult to find any research into the lexical divergence which is not part of a glottopolitical agenda which aims at either underlining or marginalising the differences.

Of Gottlieb's five arguments *for* replicating lexical matter, we see that some (Fekete, Klajn), but not all Serbian linguists (Lj. Popović) believe that replication of lexical matter *provides lexical enrichment*. Among the rest of the Serb and the Croat linguists replication of lexical matter is generally frowned upon and domestic substitutions are preferred.

We will now turn to the 'mechanics' of how replicated matter is adapted into the replica language (RL) followed by how normativists categorise replicated matter.

#### **4.2 Adaptation of replicated matter**

In research on matter replication (lexical or grammatical), a distinction is made as to the degree of adaptation to the RL. All replicated matter is more or less adapted to the RL and are at different stages on a continuum along the line from the pole 'foreign' towards the pole 'domestic'. (Bugar-ski 1996: 17) The inevitable *adaptation* of the foreign matter is what turns the matter into replica matter and not a duplicate of the model matter.

##### *Phonological and accentual adaptation (transphonemization)*

One aspect of the adaptation is phonological. Most matter in a model language (ML) simply is not pronounceable in the RL, so the pronunciation of the replica is different than the pronunciation of the model. Rudolf Filipović (1990) terms this *transphonemization* which in essence is a substitution of the pronunciation used in the ML with a pronunciation in accordance with the regularities of the phonology of the RLs – Croatian and Serbian.

Related to the phonological adaptation there is also an accentual adaptation. The accent in standard Croatian and Serbian is never on the last syllable. The accent may be long or short (quantity) and at the same time either rising or falling (quality). So, when matter is replicated into standard Croatian or Serbian, it is adapted to the standard accentual system as well as the phonological characteristics of these languages, for instance:

- a) English 'budget' /'bʌdʒɪt/ to Serbian and Croatian /'budʒet/ - *bùdžet* (short rising accent)<sup>171</sup>
- b) English 'tank' /'tæŋk/ to Serbian and Croatian /'tɛŋk/ - *tènk* (short falling accent)
- c) German 'Mundstück' /'muntʃtʏk/ to Serbian and Croatian /'muntʃtik/ - *mùndštik* (short rising accent)
- d) French 'niveau' /ni'vo/ to Serbian and Croatian /'nivo:/ - *nìvō* (short rising accent on /i/, long stress-less vowel /o:/)
- e) English '(The) Beatles' /'bi:tlz/ to Serbian and Croatian /'bitlzi/ - *Bìtlzi* (short falling accent on /i/ and a plural ending /i/)

*Morphological adaptation (transmorphemization)*

A second aspect of matter replication is morphological adaptation. Bound morphemes (affixes), i.e. grammatical matter, are rarely replicated on its own. If grammatical matter is replicated, it is replicated together with a free morpheme (a root, a lexeme). The replicated free morpheme may take on bound morphemes from the RL. The alteration in the form of the entire replicated lexeme or the addition of derivational morphemes on replicated root-morphemes is also termed *transmorphemization* by Filipović (1990: 31), for instance in:

- a) farm > *farma*. The English noun 'farm' is morphologically adapted through the addition of the suffix *-a*
- b) compatible > *kompatibil(a)n*. The adjective 'compatible' is morphologically adapted through the addition of the suffix *-(a)n*<sup>172</sup>
- c) (to) train > *trenirati*. The English verb '(to) train' is morphologically adapted through the addition of the suffix *-irati*<sup>173</sup>

There is a noticeable difference between the results of morphological adaptation of lexical matter between Serbian and Croatian which stems from the fact that similar international matter has been replicated form

---

<sup>171</sup> The short quantity combined with the rising tonality is marked by the diacritic ` above the accented vowel (or syllabic r)

<sup>172</sup> The adjective 'compatible's root is phonologically adapted from /kəm'pætɪbl/ to /'kɔmpatibi:l/

<sup>173</sup> The verb '(to) train's root is phonologically adapted from /'treɪn/ to /trɛn/

different model languages or simply because there was no common standardisation of Serbian and Croatian at the time of replication (Nyomárkay 1981: 214-216).

<u>Serbian</u>	<u>Croatian</u>	
finansijski	financijski	'financial'
demokratija	demokracija	'democracy'
minut	minuta	'minute'
leksema	leksem	'lexeme'
kriterijum	kriterij	'criterion'
lingvista	lingvist	'linguist'
hemija	kemija	'chemistry'
manipulisati	manipulirati	'to manipulate'
organizovati	organizirati	'to organise'
metod	metoda	'method'

See also subsection 1.2.2 above.

#### *Semantic alteration*

A third aspect does not relate to the adaptation of the form or the sound but of the semantic content of the replicated matter. Filipović tells us that it is the rule rather than the exception that replicated matter, i.e. the replica, does not carry the same number of meanings as the model does. Usually, the replica only has one specific meaning out of several meanings present in the model (Filipović 1990: 38).

When replicated matter has exactly the same semantic content as the model matter, it is termed *zero semantic extension*. When the replicated matter has fewer referents/meanings, it is termed *semantic narrowing*. When the replica matter takes on more meanings than the model matter, it is termed *semantic broadening* or *semantic extension*. Filipović's definition is partly equivalent to Andersen's definition of *semantic change*, in which he uses the term relexicalisation for semantic extension but he does not deal with semantic narrowing as he opposes relexicalisation to delexicalisation which covers the move from lexical towards grammatical (cf. p.121)

a) *zero extension*: The replica of the English noun 'paperback' *pejperbek* has the same semantic content as the model matter.

b) *semantic narrowing*: The replica of the English model matter 'tractor' has fewer meanings than the model. According to the "Full Definition of

TRACTOR” in *The Free Meriam-Webster’s Dictionary* ‘tractor’<sup>174</sup> has two major meanings and specifications of one of these major meanings:

1, a) a 4-wheeled or tracklaying automotive vehicle used especially for drawing farm equipment, b) a smaller 2-wheeled apparatus controlled through handlebars by a walking operator, c) an automotive vehicle with a short chassis equipped with a swivel for attaching a trailer and used especially for the highway hauling of freight; *also* : a truck with attached trailer  
2) an airplane having the propeller forward of the main supporting surfaces (Merriam-Webster Dictionary 2015)

In *Rječnik hrvatskoga jezika* (‘Dictionary of the Croatian Language’) the replica of ‘tractor’ *traktor* is defined<sup>175</sup>:

“vozilo koje služi za vuču oruđa u poljogospodarstvu, prikolica i sl.: ~ *gusjeničar*, ~ *s kotačima*, laki ~, teški ~”

(vehicle which serves to pull implements in agriculture, wagons and similar: *crawler tractor*, *wheel tractor*, *light tractor*, *heavy tractor*) (Šonje 2000: 1266)

The fact remains that *traktor* is a replica of ‘tractor’ and this replica is a phonological, morphological and semantic adaption of the model ‘tractor’. *Traktor* is *not* an exact duplicate in neither sound, form nor meaning in that the replica *traktor* has a narrower semantic content and is pronounced and declined differently than the model ‘tractor’.

c1) *Semantic extension*. When replicated matter has a broader semantic scope than the model, the extension of its meaning is induced by other factors which may or may not be a result of language contact. The noun *najlon* is a replica of the English noun ‘nylon’ which was presumably first replicated with a narrower meaning than in the model<sup>176</sup>, namely ‘a synthetic material of fibre as well as things made of this material’ (Filipović 1990: 203). The replica *najlon* now has a broader semantic scope as it also refers to ‘a plastic bag’.

---

<sup>174</sup> *Tractor* is not an inherited English word, but Latin in origin. However, according to Filipović it has been replicated from the English model into Croatian and Serbian. (1990: 39)

<sup>175</sup> In Filipović’s dictionary *traktor* is defined in a very similar way (1990: 265)

<sup>176</sup> A not replicated meaning of *nylon* (in plural) “clothing for women made of nylon that fits closely over the feet and legs and goes up to the waist” (Merriam-Webster Dictionary 2015)

c2) The semantic extension may be a result of semantic replication<sup>177</sup>, i.e. the new meaning/the new referent of the same lexical matter is in itself a result of replication. The replica *album*, presumably originally replicated from Latin, at a later stage, took on the meaning of a long-play record, replicated from English. Before that *album* had only referred to a) a book/binder containing a collection of photographs or stamps and b) a collection of drawings, illustrations, and graphics by one author (M. Popović 1997: 349)<sup>178</sup>.

The three kinds of adaption (a-c) are equally valid in Serbian and Croatian. The difference between the two standard languages and their norms lies rather in the adoption, actualization and implementation and subsequent promotion of or objection to replicated matter and less so in the adaptation of the adopted, actualized and implemented matter.

#### *Orthographic adaptation*

A fourth aspect of adaptation of replicated matter is orthographic adaptation. Orthographic adaptation is not an adaptation of the linguistic means themselves but rather a conventionalised way of rendering adapted matter in writing. In other words it has nothing to do with the linguistic system itself, but has to do with the standardised norm of writing. In this aspect there is a clear difference between the two standards. In standard Serbian, which has two standardised alphabets, the Cyrillic and the Latin alphabet, all replicated matter is adapted orthographically according to phonetic principles so that it may be rendered in the Cyrillic. The rules of orthographic adaptation are identical for both alphabets. However, when it comes to foreign names, it is, in standard Serbian when using the Latin alphabet, also possible to write names according the orthographic rules of the model language. (Klajn 2011: 115)

---

<sup>177</sup> other terms: semantic borrowing, semantic calquing, polysemious copying, polysemious semantic borrowing, loan meaning extension

<sup>178</sup> the registered meanings of *album* vary depending on the consulted dictionaries, but the LP-meaning is not present in the authoritative 1960 dictionary, published by both Matica Srpska and Matica Hrvatska. In the LP-meaning *album* occurs as early as 1982 in Klaić' dictionary of foreign words (Rječnik stranih riječi), but in Šonje's Dictionary of the Croatian language (Rječnik hrvatskoga jezika (2000)) it is limited to meaning „two or more gramophone records packed together in a carton wrapping“

*Orthographic adaptation of 'Janet' in three ways:*

a1) 'Janet' > *Dženet*

a2) 'Janet' > *Џенеџ*

b) 'Janet' > Janet

This possibility is also valid for some toponyms, e.g. 'New York' may be rendered: *Njujork*, *Нјујорк* and *New York*.

In standard Croatian the rules concerning orthographic adaptation are somewhat more complex as language users are recommended to substitute all matter which is considered to be foreign with domestic matter, and when no domestic substitute exists, the matter should be marked as foreign (Filipović 2004: 232). This rather complex situation is illustrated by the equally complex rules for dealing with replicated matter which follows in the next section.

#### **4.2.1 Normative attitudes towards replicated matter**

In the „Hrvatski jezični savjetnik“, (Croatian language guide), abbreviated HJS published in 1999 by the *Institute of Croatian Language and Linguistics*, replicated lexical matter or foreign words are categorised according to their adaptation into the Croatian (standard) language and are at the same time deemed more or less acceptable, i.e. the more adapted – the more acceptable. Thus four categories of replicated lexical matter are defined as opposed to the less discrete differentiation between 'foreign words' (*tuđe* or *strane riječi*) and 'loanwords' (*posuđenice* or *pozajmljenice*):

*Tuđe riječi/strane riječi* – 'foreign words' – acceptable only if used for stylistic reasons. They *are not* to be regarded as Croatian words. Foreign names - *strana imena* - (personal and geographic) represent an exception to this recommendation because HJS (1999: 290) says that foreign names from languages written in the Latin script should not be altered (i.e. be submitted to orthographic adaptation) and that they should be pronounced (i.e. phonologically adapted) according to the rules of transliteration prescribed in language guides (i.e. normative works) (Barić et al. 1999: 290).

*Tuđice* – 'partially adapted foreign words' are foreign words which have been orthographically adapted to the Croatian language but also contain some (for Croatian) uncharacteristic features as e.g. phonetic strings such as *mpj* (e.g. *kompjutor* > computer), or word-final consonant clusters as e.g. *kl, bl, jt, ft, jl* (in e.g. *bicikl* < bicycle, *dabl* < double, *lift* < lift, *fajl*

< file). If the author is unable to find a domestic substitute for these and other *tudice*, s/he should write them in italics, thereby indicating that they are not Croatian but foreign.

*Prilagodbenice* – ‘adaptees’ are foreign words which have been adapted phonologically (including accent) and in declensions (morphologically). Even though they are adapted in these aspects, the HJS advises against their use and recommends a substitution with *domestic* words. Under the same heading, HJS also points out that the grammatical matter (derivational affixes) used to adapt foreign verbs are *-a-*, *-ova-* and *-ira-*, of which one, *-ira-*, is itself a piece of replicated (grammatical) matter.

*Usvojenice* – ‘adoptees’ are words of foreign origin which have been completely adopted into Croatian and HJS does not recommend searching for substitutions of these words, partly because most native speakers do not consider these words foreign at all.

(Barić et al. 1999: 283–85, 290)

Clearly the normativists behind this publication intend to demote replicated matter and promote what they term domestic matter. In the event that a writer is compelled to write an unadapted foreign word, this should be clearly marked in italics, though it is not recommended to italicize foreign names. Characteristically, replication of semantics, i.e. beyond the form (phonetic, morphological and orthographic) is not discussed and there is also an emphasis on the written norm, characteristic of normative and puristic publications. The linkage, between what is acceptable and what is heritage (domestic and/or inherited), is clear in the different attitude towards replicated matter from on the one side Latin and Classical Greek as opposed to other model languages on the other. It is also revealing to compare the two different treatments of word-final consonant clusters in Petrović (1996), (cf. p.143) on Serbian phonological developments and here in Barić et al. (1999) Where Petrović observes and describes, Barić et al. opposes and proscribes. Furthermore, the distinction between the two latter categories: ‘adaptees’ and ‘adoptees’ is purely social. From a linguistic point of view they ‘fit’ equally well into the Croatian linguistic system, but nonetheless they are seen as more or less acceptable depending on extra-linguistic factors, described in the literature on language planning as ‘acceptance’, cf. above, p. 104ff.

As we saw in the examples of puristic attitudes in the linguistic advice from the state-funded Institute of Croatian Language and Linguistics

on p. 83, replicated lexical matter, termed *internationalisms* are more acceptable than other replicated matter. HJS explains that apart from the fact that words, primarily of Greek and Latin origin, are more easily integrated (i.e. adapted) into the Croatian language, 'Latin and Greek are the foundation on which our and the European culture and civilisation has been built'<sup>179</sup>. This is a very clear example of how normativists impose their view of what is good and acceptable and what is not, and besides, it is a very salient example of the Croatian national narrative which emphasises the linkage to Central Europe and not the Balkans.

In Serbia, normativist publications such as *Pravopis srpskoga jezika* ('Orthographic manual of Serbian')(Pešikan, Jerković and Pižurica 2010) *Srpski jezički priručnik* ('Language handbook of Serbian') (P. Ivić et al. 2011),<sup>180</sup> *Normativna gramatika srpskog jezika* ('Normative grammar of Serbian') (Piper and Klajn 2013)<sup>181</sup> and *Tvorba imenica u srpskom jeziku* ('Noun formation in Serbian') (Ćorić 2008)<sup>182</sup> do not object to or hierarchize replicated matter in the same manner as does the Croatian HJS. They describe and prescribe how foreign matter, especially names should be transcribed, i.e. adapted and subsequently explain the mechanisms by which replicated suffixes are used and integrated into Serbian. If we, however, turn to the decisions made and published by "The Board of Standardisation of the Serbian language" we find a more prescriptive though moderate attitude which might be best illustrated by the following quote:

'These general remarks illustrate how complicated the question of foreign words is. The general conclusion which may be drawn from them is that a foreign term should be translated: when and where it is possible or sensible.' (Brborić et al 2006: 102)<sup>183</sup>.

However, in other and more specific matters the members of the

---

<sup>179</sup> Original: „latinski i grčki [su] temelj na kojemu je izgrađena naša i europska kultura i civilizacija“ (Barić et al. 1999: 285)

<sup>180</sup> Whose first edition was published in 1991 under a slightly different name and in 2003 up-dated and extended.

<sup>181</sup> Recently published as a project organised by The Board for Standardisation of the Serbian Language

<sup>182</sup> Published by *Društvo za srpski jezik i književnost Srbije* ('The Association for Serbian language and the literature of Serbia') which as one of its main objectives sees the education of language professors in secondary school as well as development of educational and scientific literature concerning Serbian language and Culture (Društvo za srpski jezik i književnost Srbije 2015)

<sup>183</sup> Original: "Ове опште напомене само указују на то колико је питање употребе страних речи компликовано. Начелни закључак који из њих следи јесте да страни термин треба превести – тамо где је то могуће или разумно."



Board are less moderate, when they e.g. address the question of so-called 'loanwords of laziness' (*Pozamljenice iz lenjosti*) and recommend using domestic words such as *uprava* instead of *menadžment* ('management') (Ibid: 105).

So in conclusion, the normative attitudes of the leading and government-funded linguists do differ between Serbia and Croatia which well mirrors the general descriptions of Croatian and Serbian's openness towards foreign influence as well as the results of recent research outlined above, p. 146.

As we have just seen, the normative bodies in both countries prefer domestic, inherited or international matter to other foreign matter. Now we will look into the strategies applied when substitutions for the undesired replicates are to be found.

#### **4.2.2 Substitution of replicated matter**

The motivations for replicating a lexeme range from the need to name a new phenomenon to the need to express personal or group identity (Gottlieb 2006: 196). So, if language planners (individuals, pressure groups or governmental agencies) through corpus planning of the standard language aspire to counter the replication of lexical matter along with the phenomena it denotes, they need to provide alternatives to the replicated matter. This is done by substituting the replicated/foreign matter with, as the purists put it, domestic matter which is in accordance with the 'spirit' of the language.

The replicated matter is either substituted by a replicated lexical or semantic pattern, which, in Haspelmath's terminology, is done through calquing and loan meaning extension<sup>184</sup> (Haspelmath 2009: 39).

Examples of loan translation (calquing) are mostly seen in compounds, where the elements in the compound are translated element by element (cf. p. 171ff), to a lesser extent in translation of derivational affixes (cf. p. 220ff) but also in nouns (Ibid.)

---

<sup>184</sup> Calquing is also called loan translation and translation loan. Loan meaning extension is also called semantic extension, polysemious copying, semantic loan, loan-meaning and loan-shift (Muhvić-Dimanovski 1992: 103; Romaine 2004: 51).

### *Loan translation*

- 1a) Compounds: *nana čaj* (< 'mint tea'),  
*vodopad* < Ger. Wasserfall ('waterfall')  
[vod(a) – 'water', -o- (linking element), pad – '(a) fall']  
*samoposluga* < self-service  
[sam – 'self', -o- (linking element), posluga – 'service']
- 1b) Affix: *višenacionalan* (< 'multinational') [*više* – 'more']
- 1c) Noun: *računar* (Ser) / *računalo* (Cro) (< 'computer'),  
[*računati* – 'to compute, count, calculate']

According to Turk and Opašić (2008), substitutes introduced in Croatian as alternatives to replicated matter may be received in three different ways depending on whether or in which role this innovation is put into use by the speakers (Turk and Opašić 2008: 84). In other words, it is, in language planning terms, a question of *acceptance* and *implementation* because the mere official acceptance of an innovation within the standard language is not the same as its implementation into the usage norm. The substitute is either (A) implemented by the language users as was the case with the substitute: *samoposluga* (<self-service) (B) both the replicated matter and its substitute are implemented: *kompjutor* - *računalo* (<computer) or (C) the accepted substitute is not implemented by the language users: *hardver* - *\*železarija* (<hardware), *marketing* - *\*tržišтво* (<marketing) (Turk and Opašić 2008: 84).

### *Loan meaning extension*

Examples of loan meaning extension is seen in all matter, for instance:

2) mouse: a small mobile manual device that controls movement of the cursor and selection of functions on a computer display<sup>185</sup> > miš: ulazna jedinica jedinica računalnog sustava kojom se mogu unositi podatci u računalo ili odabrati naredba što će je računalo izvesti<sup>186</sup>

The polysemy of the English noun 'mouse' is extended to the Croatian and Serbian noun *miš*, that is, the polysemy of the English word 'mouse' is replicated into Croatian and Serbian, and thus the meaning „a small

---

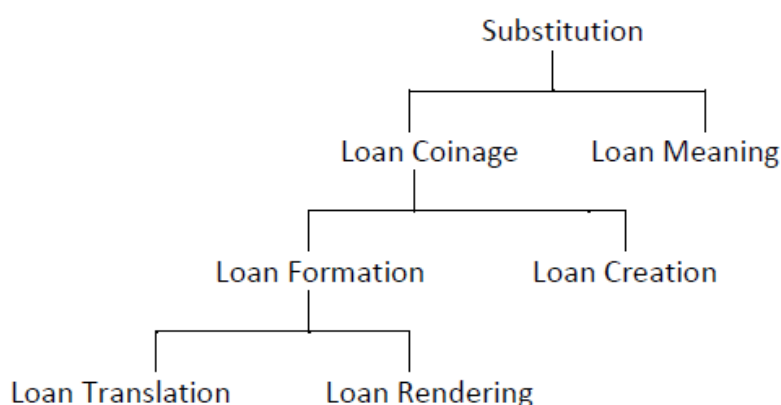
<sup>185</sup> English definition accessed (June 3, 2015) at:

<http://www.merriam-webster.com/dictionary/mouse> (June 3, 2015)

<sup>186</sup> Croatian definition in: (Šonje 2000: 598)

mobile manual device that controls movement of the cursor and selection of functions on a computer display” is extended from English to existent meanings of *miš* in Croatian and Serbian. Loan meaning extension lies within the scope of semantic alteration (as described by Filipović, cf. p. 157, above). However, in semantic alteration it is the meaning of the replicated matter, which undergoes extension, zero extension or even narrowing, whereas loan meaning extension pertains only to extension of meaning from a model to a replica language. Two additional categories of substituting replicated matter, besides calquing and loan meaning extension, has been proposed by David Duckworth (1977) who, referring to the pioneers in language contact studies, Uriel Weinreich (1953) and Einar Haugen (1950) under the umbrella term *substitution* subsumes Loan Meaning (loan meaning extension), Loan Coinage, Loan Formation, Loan Creation, Loan Translation and Loan Rendering.

Figure 4.1 - Duckworth’s categorisation of matter substitution



Duckworth (1977: 40)

Loan coinage is taken to mean the instantiation of a new word, whereas Loan meaning (extension) is the extension of meaning to an existent word, in the same way as Haspelmath (2009) describes it. A new word can be instantiated either by creating a new word (Loan Creation) to replicate the meaning of a foreign word without replicating any other aspects of the foreign word (Duckworth 1977: 52)<sup>187</sup>, as in (3), below, or by forming a new word (Loan Formation) either by a true-to-the-model element by element translation (Loan translation/ calque) as we saw in

---

<sup>187</sup> Original: “Ein Wort wird neu geschaffen, um ein fremdes Wort wiederzugeben, ohne sich formal in irgendeiner Weise an dieses fremdes Wort anzulehnen”

(1a-c), above, or by an approximate translation (Loan rendering) where the model is evident but not translated directly as in (4), below:

(3) *brz(j)av* [*brz(o)* – ‘quick’, -o- (linking element) *jav(iti)* – ‘to report’] <‘telegraph’

(4) *poveznica* [*povez(ati)* – ‘to tie, link, join, bind (up)’ + nominalising suffix -*nica*] < ‘link’ (IT-term)<sup>188</sup>

A fifth way of substituting replicated matter is revivification of archaisms or promotion of regionalisms, which are manifestations of archaic and ethnographic purism (cf. “2.2.3 Purism” above). In such instances as in (5), there is no innovation within the lexis. It is, rather, an innovation of the usage norm, and only possibly an innovation of the content, that is if the promoted regionalisms or revived archaisms take on new meaning, in which case they are instances of loan meaning extension.

(5) *časnik* (1991 - ) < *oficir* (1945-1991) < *časnik* ( - 1944)

*časnik* ('mil. officer') was, as mentioned in „1. Introduction“ (p. 29) one of the distinctly Croatian words or expressions which were suppressed during the Communist era, only to be revived in Croatian usage norm in connection with Croatia's exit from Yugoslavia. As *časnik* does not carry any new semantic content, it is only the usage norm of Croatian and not Croatian lexis, which has been innovated. So, contact-induced innovation of the lexis through the substitution or avoidance of replicated matter is done in four different ways: Loan meaning extension, Loan creation, Loan translation (Calque) and Loan rendering. However, substitution may also be achieved without innovations of the lexis.

Interestingly, in Croatian standardised usage not only replicated matter and Serbian matter, but even domestic matter was substituted, presumably, because it was identical or near-identical with Serbian domestic matter. Pranjković (2000: 72) reports the following usage changes to have happened in Croatian after the independence in 1991.

---

<sup>188</sup> ‘link’ is otherwise translated into: *karika*, *zglobo*, *veza*, *spoj* (Bujas 2008)

<u>Serbian</u>	<u>Croatian usage change</u>
<i>izvještaj</i>	<i>izvještaj &gt; izvješće</i> ('report')
<i>prilog</i>	<i>prilog &gt; privitak</i> ('attachment')
<i>pažnja</i>	<i>pažnja &gt; pozornost</i> ('attention')
<i>saopštenje</i>	<i>saopćenje &gt; prićenje</i> ('announcement')

Many other lexemes were substituted due to their linkage to Serbian either by being true serbisms as e.g. *spisak > popis* ('list'), *uputstvo > uputa* ('instruction') or by being replicated matter from other languages, also used in Serbian as e.g. *opozicija > oporba* ('opposition'), *kandidat > pristupnik* ('candidate') *centar > središte* ('centre').

The substitution of these and many other lexemes supports the notion that changes in the usage of lexis can be motivated by speakers' need to signal that they belong to a specific group (Croats) and not to another specific group (Serbs, Yugoslavs) or in Gottlieb's words: "the need to express personal or group identity" and shows that the symbolic, identity-related function of a language is so influential that it does indeed overrule the communicative function as speakers will have to learn and implement the lexis in their language usage in order to express their group identity. So, the symbolic function of language is, as discussed in section 2.2 above, crucial when language is used as a builder or protector of identity.

### 4.3 Nouns

#### *Replicated lexical matter (Importations)*

Among the parts of speech (word classes) nouns are statistically the most replicated (Matras 2009: 167). In the previously mentioned study on “English in Europe” approx. 80 per cent of the loanwords are nouns (Görlach 2004: 7) (Cf. also subsection 3.3.2, above).

The high percentage vis-à-vis other parts of speech such as verbs, adjectives, adverbs and prepositions is due to the high number of referential functions covered by nouns (Matras 2009: 168). Most of the replicated nouns refer to new concepts, things and inventions (Görlach 2004: 7). Another, rather evident reason for the higher percentage of replicated nouns, I believe, is the relatively higher percentage of nouns in languages in general. It has proven difficult to find any precise data on the typical proportional difference between nouns and other parts of speech but nouns are usually designated as the most numerous parts of speech, i.e. there generally exist more nouns than verbs, adjectives and adverbs in a language.<sup>189</sup>

Replicated acronyms as e.g. *NATO* and *SIDA* (< Fr. *SIDA*) [‘AIDS’] or eponyms as *vat* (< Watt) are in this study also considered to be replicated nouns, since they, despite their acronymic or eponymic status in the model language, are replicated as nouns, and therefore must be seen as nouns in the replica language. (Košutar and Tafra 2009: 105). The acronyms, replicated as nouns, have, as do all nouns in Serbian and Croatian, an assigned gender. Acronyms ending in /-a/ are assigned the feminine gender and enter the nominal declension for feminine and masculine nouns ending in -a<sup>190</sup>. All other acronymic nouns in both Serbian and Croatian are assigned the masculine gender and are, if declined, declined according to the declension for masculine nouns ending in /-ø/ and /-o/or/-e/<sup>191</sup> (Pervaz 1996: 109), (Babić and Moguš 2011: 75). Pervaz observes that, in the usage norm, sometimes acronymic nouns are declined, sometimes they are not (Ibid: 111). Pervaz’s observation is confirmed by the following findings in the Croatian and Serbian text corpora (HNK and SrpKor).

---

<sup>189</sup> Only the open/productive word classes are relevant in this comparison.

<sup>190</sup> All nouns ending in -e in the genitive singular belong to this declension and may be of either masculine or feminine gender.

<sup>191</sup> Masculine nouns ending in -a in the genitive singular belong to this declension

Undeclined:

[ona je] osudila agresiju NATO –

‘[she] condemned the NATO aggression’ (SrpKor)

snage NATO –

‘NATO forces (HNK)’

Declined:

[oni su] osudili agresiju NATO-a –

‘[they] condemned the NATO aggression’ (SrpKor)

snage NATO-a – NATO forces (HNK)

Both options are acceptable according to the prescribed norm (i.e. the orthographic rules) of the Serbian standard language as per 2010 (Pešikan et al. 2010: 159). According to the Croatian Standard, acronyms must be written with desinences<sup>192</sup>, written in lower case letters and must be attached through a hyphen on all acronyms written in capital letters, e.g.: *u SAD-u* – (‘in USA’), *u INA-i* (‘in INA’)<sup>193</sup>. (Jozić et al. 2013: 80)<sup>194</sup>.

Replicated lexical matter is traditionally called simply ‘loanwords’ and the replication process ‘borrowing’. Haugen distinguishes between *importation*, *partial substitution* and *substitution* (Haugen 1950: 214). As indicated in the previous section, the different degrees of adaptation into the RL have made it necessary to apply different labels in order to distinguish between the different degrees of adaptation. In German and the Nordic languages a distinction between *Fremdwörter* (foreign words) and *Lehnwörter* (loanwords) is made, the latter being the more adapted. Gottlieb (2006) proposes another distinction: *Overt* vs. *covert* lexical borrowings, an opposition dependent on the degree of adaption and the subsequent degree of transparency in the eyes of the language user (Gottlieb 2006: 199). Serbian and Croatian examples of this opposition would be ‘covert’ *keks* (<Engl. ‘cakes’), opposed to ‘overt’: *marketing* (<Eng. ‘marketing’).

As indicated throughout this thesis, I assume that covert borrowings or rather: covert influence from other languages, will precisely because of its covertness, its domestic guise if you will, be more readily accepted (and in some cases promoted) by the language authorities and

---

<sup>192</sup> With certain specific exceptions

<sup>193</sup> *INA* – *Industrija nafte* (‘Oil industry’)

<sup>194</sup> The Croatian orthographic manual of Babić and Moguš (2011) prescribes otherwise (Babić and Moguš 2011: 75–76)

also presumably be more readily implemented by the language users. The focus on overt lexical borrowings (replicated lexical matter) has been and is still very large among scholars and laymen in the Serbian and even more so in the Croatian speech community. The Croatian linguistic journal *Jezik* has in Turk and Opašić's words: 'constantly been publishing articles on Anglicisms and their status in the Croatian lexis.' (Turk and Opašić 2008: 83). The Serbian linguist Branislav Brborić similarly concludes that *Jezik* has been preoccupied with the issue of loanwords (*tudice*) from it was established in 1952 and, at least, until 1991 whereas its Serbian counterpart *Naš jezik*, established in 1950, has very few articles on the theme (B. Brborić 1996: 30). Objections to loanwords do, however, exist among Serbs and also among Serbian linguists as we have seen in the recommendations from *The Board of Standardisation of the Serbian language*.

### ***Substitutions***

The motivations for replicating a lexeme range from the need to name a new phenomenon to the need to express personal or group identity by preferring a foreign or native expression for a given lexical content. So, if a speech community or the linguistic authorities aspire to counter the replication of lexical matter along with the new phenomena it denotes, substitutions have to be found or created. New nouns in a language will, very rarely, have been created *ex nihilo* as a *root creation*, also termed *coinage*. In English the noun 'hobbit' is considered to be such an unmotivated creation, cf. also p. 124. Other nouns that have been created without any linkage to other linguistic matter are the onomatopoeia as for instance, animal noises, which are believed to be imitations of non-linguistic sounds: *meow*, *cock-a-doodle-do*, *oink*, etc. All other wordforms are either derived from existing inherited linguistic matter or enter the language through a replication of linguistic matter.

There are, as we saw in the section on grammatical and lexical change, p. 118ff several ways in which a noun (or any other wordform) may be formed intralinguistically: *derivation*, *conversion*, *clipping*, *blending*, *acronyms* and *compounding*. In this section the focus is primarily on the *replicated* nouns (incl. acronyms), but as we are also looking into the accepted and/or implemented alternatives to some of the replicated nouns, it is worth noting that the proposed and more or less successfully promoted newly formed substitutions, are mostly formed through derivation (sometimes calqued) and compounding (mostly calqued).



The proposed substitutions are, however, not always new inventions, but may also be revived archaisms or loan meaning extensions of existing inherited vocabulary, here shown with Croatian examples.

<u>Replicated noun</u>	<u>Substitute</u>
<i>kompjutor</i> (< computer)	<i>računalo</i> (calque, derived from the verb <i>računati</i> – to compute)
<i>muzika</i> (< Gr. musikē) [music]]	<i>glazba</i> (revivified, derived from the noun <i>glas</i> – sound)
<i>tajm-aut</i> (< time out (in sports))	<i>predah</i> (loan meaning extension - a breathing spell, a breather)
<i>vešmašina</i> (< Ger. Waschmaschine)	<i>perilica</i> (derived from the verbal root <i>per-</i> (wash)-[washing machine])
<i>hardver</i> (< hardware) the noun <i>željezo</i> (iron))	<i>željezarija</i> (loan creation, derived from the noun <i>željezo</i> (iron))
<i>oficir</i> (< Fr. officier)	<i>časnik</i> (revivified, after having been banned in Socialist Yugoslavia)
<i>kontakt grupa</i> <sub>CMP</sub> (< contact group)	<i>doticajna skupina</i> <sub>NP</sub> (calque: <i>doticajni</i> , adj. (contact), <i>skupina</i> , n. (group))
<i>marketing</i> (< marketing)	<i>tržišтво</i> (loan rendering, derived from <i>tržnica</i> (market-place) or <i>tržnik</i> (marketing professional))
OSCE	(Serb.) <i>OEBS</i> / (Cro.) <i>OESS</i> (translated abbreviations)
OSCE - (Organisation for Security and Co-operation in Europe)	<i>OEBS</i> – ( <i>Organizacija evropske bezbednosti i saradnje</i> ) <i>OESS</i> – ( <i>Organizacija europske sigurnosti i suradnje</i> )

The focus on loanwords in the Croatian periodical *Jezik* is still present. Every year since 2006 (and also in 1993, 1994 and 1998), *Jezik* has announced a competition in which the editors ask their readers to come up with new Croatian words to substitute loanwords, primarily ‘unnecessary English ones’ (Ham 2011a: 73–74). Every year the goal is to find

three winning words. Some years (2009 and 2012) the candidate words have not met the criteria, but in the years 2006-2012 the following substitutions for loanwords have been found recommendable.

2006: 1) *uspornik* ('speed bump') < *ležeći policajac* < Ger. *liegender Polizist* or Eng. *sleeping policeman* - derived from *usporiti* - 'to slow down' (tr.)

2006: 2) *smećnjak* ('garbage container/dumpster') < *kontejner za smeće* ('container for garbage') - Derived from: *smeće* - 'garbage' (Ham 2007: 120)

2007: 1) *opuštaonica* ('indoor relaxation place') < *wellness* - Derived from: *opuštati se* - 'to relax' (Ham 2008: 107)

2008: 1) *proširnica* < *žilni potporanj* ('stent') - Derived from: *proširiti* - 'to widen' (Ham 2009: 74)

2010: 1) *ispraznica* < *floskula* < Ger. *Floskel* ('empty phrase, platitude') - Derived from: *isprazniti* - 'to empty'

2010: 2) *osjećajnik* < *emotikon* < 'emoticon' - Derived from: *osjećati (se)* - 'to feel' or *osjećaj* - 'feeling, emotion'

2010: 4) *svidalica* < a 'like' (on Facebook) - Derived from *svidati se* - 'to like', 'to be fond of' (Ham 2011a: 73)

2011: 1) *zatipak* < *tipfeler* < Ger. *Tippfehler* ('typing error') - Derived from *tipkati* - 'to type'

2011: 3) *dodirnik* < touch screen - Derived from *dodirnuti* - to touch (Kekez 2012: 80)

Only the winning words which are substitutions of foreign lexis are mentioned here and what is noticeable is that *all* the winning words are nouns.

#### 4.3.1 Empirical data

Having made parallel searches of the Croatian National Corpus, v. 2.5 (HNK) and the Corpus of Contemporary Serbian version 2013 (SrpKor) I can confirm that Serbian language usage indeed does contain a more frequent use of replicated lexical matter than does the Croatian. Similarly, it is clear that substitutions for internationalisms are used more frequently in Croatian than in Serbian. Substitutions for Anglicisms, when existing in both languages, are however used less frequently.

I have searched the two corpora for 61 different R-nouns (replicated nouns) and their substitutions in Croatian and Serbian. I have selected these nouns from several different articles on lexical innovations considered to have increased in usage since the 1990s. The selected

nouns are listed in Pranjković (2000), Turk and Opašić(2008), Dešić (2009) and Vasyľeva (2009). In addition, I have added a few from my personal experience in navigating the vocabularies of Serbian and Croatian. For summaries and discussions of these and other articles, see section 4.1 above. Some of the nouns are in the literature, including grammars for foreigners (Alexander 2006; Mønnesland 2002) mentioned as examples of the lexical divergence between Croatian and Serbian. In the following tables I have marked the lexical pairs according to the original authors who have listed them: Pranjković (2000) – P, Turk and Opašić (2008) – TO, Dešić (2009) – D, Vasyľeva (2009) – V and own additions – M.

In the tables the queried lexical pairs are ordered alphabetically according to the replicated noun. In some instances there are more than one domestic equivalent, either because there is a difference between the standard Croatian and the standard Serbian equivalent as in *jamstvo* [Croatian standard] and *jemstvo* [Serbian standard] which are domestic equivalents to the internationalism *garancija* ('guarantee') or because the replicated noun is semantically broader than the Croatian or Serbian equivalents, as in *čimbenik* and *činitelj* [Croatian standard] substitutions for the internationalism *faktor* ('factor'). Among the domestic equivalents to the replicas we find examples of Loan translations, Loan meaning extensions, Loan renderings, Loan creations and revived nouns, as exemplified above, i.e. all the described means to replenish the lexicon with domestic rather than replicated matter are in play. For a detailed description and illustration of the querying method, see Appendix D, p. 333.

### ***The model languages***

A rather large number of the queried nouns are so-called internationalisms (of Classical Greek or Latin origin), which represent the kind of replicated nouns which are acceptable to the Croatian normativists, i.e. have a place in the codified standard language. They may have been replicated directly from the classic languages or via another contact language<sup>195</sup>. Their use may therefore increase as a consequence of contact with, e.g. English. Besides English (E), Classical Greek (H) and Latin (L), German (G), French (F) and Italian (I) are also represented in the sample. When a replicated noun is not inspired by an English model but also present in English (i.e. a similar replica exist in English) it is marked as “in E”. A

---

<sup>195</sup> also called a proxy-borrowing

smaller number is replicated directly from English or other languages with which Serbian and Croatian language users have been or are in contact. These replicas are, as we recall, not acceptable in the eyes of the Croatian normativists except under certain stylistic circumstances. In Tables 4.3.1-4.3.6, the replicas are marked for model language in the column “ML”.

Let us recall that the recommendation from normativists in Croatia is not to use replicated lexical matter as long as a domestic equivalent exists while the recommendation from some Serbian normativists is less straight-forward as it prescribes the use of domestic equivalents when it is both possible and sensible to do so. Other Serbian leading normativists only describe the proper way to adapt (phonologically, morphologically and orthographically) replicated lexical matter, cf. above.

The results given below are only representative for the usage norm of the Publicistic functional style because the queried corpora only contain text representing the usage norm of the Publicistic functional style, cf. p. 45ff.

#### *The symbolic value*

When the symbolic value of a new expression is higher than its communicative value (as discussed in section 2.2 above) it should have a better chance of being implemented and actualised by the language users, as envisaged by Andersen, cf. p. 116.

According to Andersen’s envisaged process of a linguistic innovation becoming part of the language, it is vital that ‘cohorts of speakers’ actualize the innovation. If we generalise this view to cover innovations introduced in the usage norm of a language, substitutes or replicas with a high usage, should have a better chance of becoming part of the usage norm.

#### *Synonymy*

The domestic equivalent/substitute need not be synonymous with the replica in all aspects. In fact, strictly speaking, synonyms do not exist. As words always appear in context and are always coloured by their typical contextual properties, there are no actual synonyms. Between any two ‘synonyms’ there will always be differences in which other words they more easily collocate with and they will often appear in different genres and text types. (Gottlieb 2014: 11) A replica may be the norm in one functional style and not in the other. It is also sometimes the case that the substitute only covers some of the meanings that the replica covers, as is

the case with the lexical pair *disciplina/stega*, where *stega* covers part of the meaning but not all the potential meanings of *disciplina*. Thus *stega* covers meaning (1a) ‘system of rules of behaviour with which order is accomplished’ and (1b) application of rules of behaviour, but not meaning (2) ‘A defined area within science or sports which has its own subject matter and methods’ (Šonje 2000: 184). The opposite situation is also seen, i.e. when the domestic equivalent covers not only the meaning(s) of the replica but has additional or a more general meaning. In the results listed in the following tables, whenever a situation of discrepancy as the *disciplina/stega*-situation above is valid and has led me to conclude that the result is misleading, I have modified the numbers manually in accordance with my analysis, i.e. reading of the individual concordances,

Bearing these caveats in mind, let us begin the analysis of the queried nouns. I have queried exactly 61 replicas and 63 substitutes. The discrepancy in numbers is due to the fact that one replica had more than one substitute, viz. *faktor* (factor), which I have paired with three different domestic equivalents: *činitelj*, *čimbenik* and *činilac*. In Table 4.3.1-4.3.6 you find all the queried replicas and their domestic equivalents in Croatian and/or Serbian along with a translation into English and the percentage “%” of each token found in both corpora. Under “Listed in” you may see the inspirational source to query these particular lexical pairs, and the column under the heading ML shows the model language for each of the replicas, and finally I have indicated under “in E” whether the replicas also exist in English.

35 or more than half of the queried lexical pairs show a clear preference<sup>196</sup> for the domestic equivalent in Croatian, whereas in Serbian the replica is either in categorical or almost categorical use<sup>197</sup>. When a replica is in categorical use, I interpret this as meaning that the substitute is not a part of the usage norm at all. In other words, when a word is only used in HNK and not in SrpKor it may be viewed as a word in exclusive Croatian use but the distribution of replica vs. domestic equivalent in HNK may still be viewed as an indication of how successful the puristic efforts of the Croatian linguistic authorities have been. By dividing the results according to how categorically the nouns are used we arrive at six types of lexical pairs.

---

<sup>196</sup> A minimum of 69 pct.

<sup>197</sup> 99-100 pct.

The first type consists of contrasting pairs, in which a replicated noun is used categorically or almost categorically in SrpKor whereas the domestic equivalent is used categorically or almost categorically in HNK.

Table 4.3.1 – Type 1 (15 pairs of nouns)

LEXICAL PAIR		Srpkor	HNK			Listed
replica / domestic equivalent	Translation	%	%	ML	In E	in
plej-of / doigravanje	<i>play-off</i>	93/07	02/98	E	+	P
penzija / mirovina	<i>pension(benefits)</i>	100/00	01/99	L	+	D
penzioner / umirovljenik	<i>retired person</i>	100/00	01/99	L	+	M
univerzitet / sveučilište	<i>university</i>	100/00	01/99	L	+	D
pasoš / putovnica	<i>passport</i>	100/00	02/98	L	+	D
firma / tvrtka	<i>company, firm</i>	100/00	03/97	L	+	V
period / razdoblje	<i>period</i>	94/06	03/97	H	+	M
muzičar / glazbenik	<i>musician</i>	100/00	04/96	L	+	M
muzika / glazba	<i>music</i>	100/00	04/96	L	+	D
funkcioner / dužnosnik	<i>official</i>	100/00	03/97	L		P
fabrika / tvornica	<i>factory</i>	99/01	00/100	L		D
portparol / glasnogovornik	<i>spokesperson</i>	98/02	00/100	F		D
nivo / razina	<i>level</i>	100/00	02/98	F		M
štampa / tisak	<i>the press</i>	100/00	03/97	I		P
golman / vratar	<i>goal keeper</i>	94/06	04/96	G		D

Since the lexical pairs in Table 4.3.1 represent a categorical split between the two corpora, they are available for symbolic signalling, i.e. for signalling belonging to this or that group, for signalling national identity.

The second type consists of lexical pairs in which the replicated noun is used categorically or almost categorically in SrpKor whereas its domestic equivalent is used mostly but *not* categorically in HNK. In the future, the lexical pairs of this type might become categorically split as are the lexical pairs of the first type but this type may also represent variable nouns in Croatian (but not in Serbian)

Table 4.3.2 - Type 2 (15 pairs of nouns)

LEXICAL PAIR		Srpkor	HNK			Listed in
replica / domestic equivalent	Translation	%	%	ML	in E	
biblioteka/knjižnica	<i>library</i>	100/00	21/79	H		D
vešmašina/perilica	<i>washing machine</i>	100/00	07/93	G		TO
kasarna /vojarna	<i>mil. barracks</i>	100/00	06/94	F		P
protest / prosvjed	<i>protest</i>	100/00	13/87	L	+	D
municija /streljivo	<i>ammunition</i>	100/00	10/90	L	+	P
opozicija / oporba	<i>opposition</i>	100/00	24/76	L	+	P
komanda /zapovjedništvo	<i>mil.headquarters</i>	98/02	23/77	I	+	P
fotokopija /preslika	<i>photocopy</i>	99/01	16/84	H+L	+	V
ekonomija /gospodarstvo	<i>economy</i>	100/00	13/87	H	+	P
ambasada/veleposlanstvo	<i>embassy</i>	99/01	09/91	F	+	D
oficir /časnik	<i>mil./pol. officer</i>	100/00	11/89	F	+	P
regrutovanje/ novačenje	<i>recruitment</i>	100/00	22/78	F	+	P
propaganda / promidžba	<i>propaganda</i>	100/00	31/69	L	+	P
pumpa / crpka	<i>pump</i>	100/00	26/74	G	+	P
grupa /skupina	<i>group</i>	99/01	28/72	F	+	D

The third type of lexical pairs is a type in which the replicated noun is used categorically or almost categorically in SrpKor whereas it is variable in HNK with either equal or wide usage as opposed to its domestic equivalent. The difference between type 2 and 3 relates only to the usage in HNK. In type 2 the domestic equivalent mostly but not categorically in use whereas in type 3, the usage is quite variable, ranging from equal usage as in *komisija/povjerenstvo* to predominant usage of the replicated noun as in *demilitarizacija/razvojačenje*.

Table 4.3.3 – Type 3 (12 pairs of nouns)

LEXICAL PAIR		Srpkor	HNK			Listed in
replica/domestic equivalent	Translation	%	%	ML	In E	
lingvistika/jezikoslovlje	<i>linguistics</i>	98/02	42/58	L	+	P
major/bojnik	<i>major</i>	99/01	45/55	L	+	P
komisija/povjerenstvo	<i>committee</i>	100/00	51/49	L	+	D
patrola/ophodnja	<i>patrole</i>	100/00	54/46	F	+	P
direktor/ravnatelj	<i>manager,director</i>	100/00	59/41	L	+	V
atmosfera/ozračje	<i>atmosphere</i>	100/00	69/31	H	+	P
kvaliteta/kakvoća	<i>quality</i>	99/01	70/30	L	+	P
centar/središte	<i>centre</i>	95/05	74/26	L	+	P
generacija/naraštaj	<i>generation</i>	94/06	74/26	L	+	P
gimnastika/tjelovježba	<i>gymnastics</i>	100/00	74/26	H	+	D
demilitarizacija/razvojačenje	<i>demilitarisation</i>	97/03	77/23	L	+	P
disciplina /stega	<i>discipline</i>	99/01	86/14	L	+	P

The fourth type consists of lexical pairs which are variable in both corpora, but with a clearly higher usage of the domestic equivalent in HNK. In SrpKor there seem to be no clear pattern but the distribution of the paired nouns clearly differs from the distribution in HNK and the usage of a replicated noun is always greater in SrpKor than in HNK. When a replicated noun (e.g. *factor*) is in large use in HNK (47%) it is in wider use in the SrpKor (76%). And when a replicated noun (*bukmejer*) is only in moderate use in Serbian (26%) it is in even less use in HNK (10%)

Table 4.3.4 – Type 4 (9 pairs of nouns)<sup>198</sup>

LEXICAL PAIR		Srpkor	HNK			Listed in
replica / domestic equivalent	Translation	%	%	ML	In E	
štimung/ugodaj	<i>atmosphere, feel</i>	49/51	09/91	G		P
bukmejer/kladioničar	<i>bookmaker</i>	26/74	10/90	E	+	V
garancija/jamstvo	<i>guarantee</i>	86/14	28/72	F	+	V
princip/načelo	<i>principle</i>	68/32	29/71	L	+	P
uniforma/odora	<i>uniform</i>	88/12	31/69	L	+	V
kompjutor/računalo	<i>computer</i>		25/75	E	+	TO
kompjuter/računar	<i>computer</i>	49/51		E	+	M
factor/činitelj or čimbenik	<i>factor</i>		47/53	L	+	M
faktor/činilac	<i>factor</i>	76/24		L	+	F

A fifth type might be conceived as a type of lexical pairs where the usage is variable in Serbian but categorical in Croatian. In the data we only find

<sup>198</sup> 7 replicas and 10 domestic equivalents



one example of this type which might be the future distribution of some of the pairs in Table 4.3.4, i.e. if the symbolic signalling of using the domestic equivalent (i.e. *računalo*) instead of the replicated noun (*kompjutor*) becomes sufficiently important to the language users.

Table 4.3.5 – Type 5 (1 pair of nouns)

LEXICAL PAIR		Srpkor	HNK			Listed in
replica / domestic equivalent	Translation	%	%	ML	In E	
sala / dvorana	hall, auditorium	57/43	02/98	G		P

The sixth and final type consists of lexical pairs where there is a categorical or almost categorical use of the replica. In other words these lexical pairs do not represent lexical pairs in either language, nor do they tell us anything about the possible differences between Serbian and Croatian usage norm in the period covered by the searched corpora. In that period, they represent non-implemented domestic equivalents to the replicas in question. The domestic equivalents could be interpreted as failed attempts at substituting replicated matter but they might also be domestic equivalents which will/have become actualised and implemented in the following period of time.

Table 4.3.6 – Type 6 (11 pairs of nouns)

LEXICAL PAIR		Srpkor	HNK			Listed in
replica / domestic equivalent	Translation	%	%	ML	In E	
epruveta/kušalica	test tube	100/00	100/00	F		P
bodybuilding/tjelogradnja	bodybuilding	100/00	100/00	E	+	V
hardver/željezarija	hardware	100/00	100/00	E	+	TO
kursor/pokazivač	cursor	100/00	100/00	E	+	V
marketing/tržništvo	marketing	100/00	100/00	E	+	TO
marš/hodnja	march	100/00	100/00	F	+	P
tajmout/predah	time-out	100/00	100/00	E	+	TO
civilizacija /uljudba	civilisation	100/00	94/06	L	+	P
hit /uspješnica	hit (i.e. success)	100/00	95/05	E	+	V
kandidat/pristupnik	candidate	100/00	99/01	L	+	P
kvantiteta/kolikoća	quantity	100/00	99/01	L	+	P

The domestic equivalents in Table 4.3.6, proposed by Croatian linguists as substitutions for the replicated nouns, were not in use or scarcely in use in the usage norm of either language as documented in the HNK and SrpKor. Six of the lexical pairs in Table 4.3.6 are replicated from English and except for *epruveta/kušalica* ('test tube') the remaining also exist as replicas in English. So, despite the recommendations of Croatian linguistic authorities these replicas are not only in use, but in exclusive use. And,

despite the successful implementation and actualisation of other substitutes in Croatian usage (Table 4.3.1-4.3.5), some replicas, which for the most part are also a part of the English vocabulary, stay in exclusive use.

Overall, the results found in SrpKor, as shown in the six tables above, indicate that a) not surprisingly the Serbian usage norm is not influenced by what is recommended use by the Croatian linguistic authorities (Table 4.3.1.-4.3.3 and 4.3.6) and b) when there is a Serbian domestic equivalent in use, it is in less than or equal use with the replica (table 4.3.4 and 4.3.5). In other words the Serbian usage norm in the publicistic functional style does not heed the recommendation to use a domestic equivalent when possible and sensible.

Generally, the findings in HNK indicate that the symbolic value of using a distinctively Croatian word is so high that not only English and German loanwords as e.g. *kompjutor*, *štimung* and *vešmašina* are less used than their Croatian equivalent but also older replicas from the French and the Classical languages as e.g. *period*, *firma*, *fabrika*, *protest*, *municija* and *ekonomija* (Table 4.3.2) etc. are being used much less than their domestic equivalents and that, generally, replicated nouns are used less in Croatian than in Serbian. The results evident in the tables show us that the proposed Croatian substitutes are in use and thus belong to the usage norm of Croatian while not being present in the Serbian usage norm. With the exception of the proposed substitutes in Table 4.3.6 they are in a rather wide use, in some cases as *dužnosnik* and *mirovina* having truly substituted the replicas *funkcioner* and *penzija* ('official and retirement', Table 4.3.1), in other cases simply coexisting with the replica as in the lexical pairs *direktor/ravnatelj* ('director') and *patrola /ophodnja* ('patrol', Table 4.3.3). So, not all of the suggested substitutes in Table 4.3.1-4.3.5 have entered the usage norm with the same force, but they have all been implemented.

#### 4.4 Nominal compounds

In this section I will, before I proceed to my analysis of empirical data, give a definition and a universal categorisation of compounds, provided by *The Oxford Handbook on Compounding* (Lieber and Štekauer 2011) and subsequently map the existing compound patterns in Serbian and Croatian. In the empirical analysis I will focus on compounds which have either one or no replicated lexical element in order to ascertain their place in the usage norms of Serbian and Croatian, resp.

Many compounds in Serbian and Croatian are believed to derive from language contact. At the turn of the 19<sup>th</sup> century we see examples of compounds in different word categories, composed of nouns, adjectives and pronouns replicated mostly from German and Czech (Zett 1969: 106) such as: *gradonačelnik*, *kolodvor*<sup>199</sup>, *sveopći* and *miomirisan* replicated from German *Bürgermeister* - 'mayor', *Bahnhof* - 'railway station', *allgemein* - 'general, universal', *wohlriechend* - 'fragrant' and *tjelovježba* replicated from Czech *tělocvik* - 'gymnastics' (Rožić 1904: 10, 15, 18, 46, 49) as well as older compounds replicated from Ancient Greek (via Church Slavonic) as e.g. *bogoslov* < Ch.Sl. *bogoslovъ* < Gr. *ὁ θεολόγος* ('theologian') and *bratučed* < Ch.Sl. *bratučedъ* < Gr. *ὁ ἀδελφόπαις* ('nephew') (Zett 1970: 162, 165).

Nominal compounds, i.e. constructions which consist of at least two content items, where the second (right-hand) element is a noun are increasing in number in both Serbian (Radovanović 2009: 210) and Croatian (Starčević 2006: 645). Many of these are clearly examples of matter replication with some orthographic, phonological and morphological adaptation. Others are combinations of replication of expression and content (*internet* < 'Internet') or solely content (*stranice* < 'pages'). In the ensuing empirical analysis I will focus upon nominal compounds where the first element consists of an undeclined noun and where there is no apparent element linking the two nouns.

The pattern: undeclined noun followed by a declined noun: [N<sub>UNDECL</sub>N]<sub>N</sub> is in wide use, not only when both elements have been replicated but also with at least one of the elements being a domestic noun as exemplified by various Serbian and Croatian linguists (domestic nouns are underlined):

---

<sup>199</sup> *kolodvor* may also be a calque of Hungarian 'pályaudvar' (Turk and Opašić 2008: 83).

*kamp prikolica* - 'caravan',  
*parking prostor* - 'parking lot' (Radovanović 2009: 210)  
*internet veza* - 'internet connection',  
*biznis klub* - 'business club' (Slijepčević 2013: 326).  
*basket parket* - 'basketball court',  
*bejzbol avantura* - 'baseball adventure',  
*spomen ploča* - 'memorial plaque' (Stanković 1997: 94–95)  
*nana čaj*<sup>200</sup> - 'mint tea' (T. Prčić 2006: 415)  
*višnja čaj* - 'cherry tea' (Horvat and Štebih-Golub 2010: 13)  
*rok zvezda* - 'rock star',  
*koka-kola novac* - 'coca-cola money',  
*pop izraz* - 'pop expression' (Čupić 1996: 314).  
*shopping kartica* - 'shopping (credit) card',  
*jagoda cocktail* - 'strawberry cocktail' (Starčević 2006: 650).  
*poklon paket* - 'gift package',  
*pop pevač* - 'pop singer',  
*čarter let* - 'chartered flight',  
 (Klajn 2011: 100–101; Piper and Klajn 2013: 250),

These compounds show that we are not only dealing with compounds replicated in their entirety from a model language but that the *pattern* is replicated as well.

The normative works and institutions within the speech communities discourage replication of the [N<sub>UNDECLN</sub>]<sub>N</sub>-pattern which is also realised in these examples, and prescribe the use of an adjectival, declinable modifier instead of a noun, e.g. \**kontakt grupa* > *kontaktna grupa* and \**internet-stranice* > *internetske stranice* (cf. also p. 83), i.e. effectively transferring the construction back to a well-known and widely used pattern of an NP, viz. [A N]<sub>NP</sub>

It is undoubtedly the existence of similar constructions that led a research project devoted to the lexical impact of English on 16 European languages to conclude that, due to the many compounds replicated from English, the word-formation patterns of Slavic but also Romance languages have begun to change (Görlach 2004: 9).

---

<sup>200</sup> *čaj* and *kartica* is considered to be domestic noun, or as the authors of HJS put it, „adoptees“ because most native speakers do not consider these words foreign at all, cf. *Usvojenice*, p.160

#### 4.4.1 The problem of categorisation

As a rule, the [N<sub>UNDECL</sub>N]<sub>N</sub> compound is not recognised as a compound in Serbian and Croatian normative works. If recognised as compounds (*složenice*) at all, they are labelled semi-compounds (*polusloženice*) (Piper and Klajn 2013: 250; Barić et al. 1995: 297). According to Klajn (2011), to distinguish between on the one side compounds and semi-compounds and on the other phrases (e.g. NPs) four basic principles exist. Klajn explains these principles as involving semantics (meaning), morphology (form), lexis and phonological stress but adds that these criteria (cf. below), however, are not sufficiently delimited.

A combination of two (or more) lexemes is a compound or a semi-compound:

1. If the combination of lexemes has **another meaning** than its individual elements (semantics)
2. If one of the elements **does not exist as an independent word** (morphology/lexis)
3. If the first (left-hand) element is **indeclinable** (morphology)
4. If it has **unique stress**, i.e. if the stress is on only one of the elements (phonology) or possibly on the linking element (Klajn 2011: 98)

The semi-compounds as e.g. *kamp-prikolica*, *internet-stranice*, *dizel-motor* and *spomen-ploča* are in the Serbian grammar by Piper and Klajn (2013: 250) defined as 'expressions of two words which constitute one meaning'<sup>201</sup> (in compliance with criterion 1). The first element is indeclinable (in compliance with criterion 3). However, both elements *do* exist as independent words (which is non-compliant with criterion 2) and 'both words retain their stress' (which is non-compliant with criterion 4), e.g. *dízel-mòtor* (/ˈdi:zel ˈmɔtɔr/). The definition of compounds and semi-compounds in the Croatian grammar by Barić et al. (1995) is similar and also distinguishes between compounds and semi-compounds using criterion 4 and it also underlines the indeclinability of the first element (criterion 3) (Ibid: 297).

Most 'genuine' compounds (those which fulfil all four criteria) have a linking element between the two component parts (an interfix), usually

---

<sup>201</sup> Original: „Полусложенице су изрази од две речи који чине јединствен појам, али свака задржава свој акценат“

an /-o-/ as in *poljoprivreda* ('agriculture'), and in some instances an /-e-/ as in *oceubica* ('patricide'). If the first component part of the compound ends in /-o/ or /-e/, this vowel is also regarded as a linking element (*drvo* ('tree') + *red* ('row') > *drvored* ('row of trees')). However, compounds without any linking element also exist (*imen* ('name') + *dan* ('day') > *imen-dan* ('name-day')) (Piper and Klajn 2013: 249). Barić (1980: 39) subcategorises nominal compounds in a similar manner, but regards the -o in *drvored* as unanalysable and the non-existent linking element in *imendan* as a 'zero linking element (-Ø-)'.

Both grammars describe compounds in various word categories (adjectives, nouns, verbs, adverbs) and categorise all compounds whose second element is a noun as 'nominal compounds' (*imeničke složenice*), i.e. a compound which is also a noun.<sup>202</sup> Both grammars also prescribe that semi-compounds are to be written hyphenated whereas full compounds are written as one word without a hyphen (Piper and Klajn 2013: 250) and (Barić et al. 1995: 297). None of the two grammars deal with unhyphenated [N<sub>UNDECL</sub>N]<sub>N</sub>, which is probably due to their focus on the standardized, prescribed norm of the languages and unhyphenated [N<sub>UNDECL</sub>N]<sub>N</sub> as e.g. *internet stranice*, *džez muzika* – 'jazz music', *velnes centar* 'wellness centre', *porno film* 'pornographic movie' are not considered part of these standardised languages, i.e. they are viewed as instances of incorrect language, i.e. unacceptable according to the prescribed norm.

However, the phenomenon under investigation in this study is often written in two words by the speakers and thus a part of the usage norm. This could be a mainly orthographic problem and a problem of unsuccessful acquisition planning; the usage norm deviates from the prescribed norm. But, it is rather, in my opinion, a problem of definition of word classes.

The scholars that do deal with these unhyphenated constructions categorise them either as NPs where the first (left-hand) element (henceforth E1) is an adjectivised undeclined nouns (D. Šipka 2010; Filipović 1990), as NPs with undeclined determiners (Slijepčević 2013; Horvat and Štebih-Golub 2010) or as NPs with a noun in E1-position as a modifier (Starčević 2006). So, in the literature, there is some disagreement as to whether constructions such as [N<sub>UNDECL</sub>N]<sub>N</sub> is best treated as an issue of

---

<sup>202</sup> Word-formation in Serbian and Croatian is, in other words, in compliance with the right-hand rule which states that it is the right-most morpheme in a word that determines the category of that word. (Štekauer and Lieber 2005: 64: 402)

word-formation, i.e. compounding, or whether it rather should be treated as a syntactic innovation where undeclined nouns modify other nouns.

Danko Šipka, who focuses on compounds with foreign undeclined elements, claims that compounds with these foreign indeclinable elements as the first element (E1) show that a noun such as *džez* or *velnes* ceases to be a noun and takes on adjectival modifying qualities. (D. Šipka 2010: 459–68). In other words; the constructions do not, in Šipka's view, represent an innovative way of forming compounds of two nouns but rather a well-known way of forming compounds where the first element is an indeclinable adjective, cf. *svetložut* 'light yellow'. Šipka's claim that [N<sub>UNDECL</sub>N]<sub>N</sub>-constructions are comparable to compounds of the *svetložut*-type is, however, easily refutable as the /-o/ in *svetlo-* is not an undeclined ending but an interfix, as exemplified by *drvored* ('line of trees') above.

Starčević, following the same logic, suggests that these constructions should be viewed as an example of syntactical innovations in Croatian induced by an extraordinarily forceful influence from English, the innovation being that nouns may be attributive modifiers in E1-position in NPs (Starčević 2006: 653).

Filipović, too, categorises the E1-positioned undeclined nouns in these constructions as adjectives:

rok      opera  
AUNDECL   N  
(Filipović 1990: 231)

Klajn has a similar point when he categorises all left hand elements as determiners be they bound morphemes or unbound undeclined nouns, adjectives or adverbs (Klajn 2011: 99).

Slijepčević uses Klajn's terminology and adds that an undeclined determiner may consist of a lexeme, which in other contexts is declinable, e.g. *internet veza* 'internet connection' and it may consist of a bound formant (*nesamostalni formant*), e.g. *akva-park* 'aqua park' (Slijepčević 2013: 322). So, Slijepčević groups the determining (or modifying) nouns (as *internet*) with lexical, but bound items as *akva* recognising that the noun *internet* and the prefix-like item *akva* fulfil the same syntactic (or word-formational) function.

Stanković, who focuses on Serbian publicistic empirical data, takes a different approach and claims that while the [N<sub>UNDECL</sub>N] by some may be

viewed as a syntactic innovation, what happens when these constructions multiply, i.e. a large number of new compounds have entered the language often as a result of replication from English, is in fact lexicalisation (word-formation) and they are thus nominal compounds in Serbian, not NPs. (Stanković 1997: 94). It is arguably so, in Stanković's view, because: it is a 'linguistic fact' that 'In the syntagmatic system of the Serbian language a construction in which a noun is added to another noun and in its indeclinable form performs an adjectival function, *does not exist*.'<sup>203</sup> (Ibid: 95)

Horvat and Štebih-Golub (2010: 9, 11) simply describe N<sub>UNDECL</sub>N-constructions in Croatian as juxtapositions, i.e. the positioning of two nouns next to each other in an unalterable order, where the first noun, which is undeclined, determines the second<sup>204</sup>, e.g.:

*sladoled kocka* ('ice cream cube')

DET        HEAD

N<sub>UNDECL</sub>    N

I believe that the lack of unity in the descriptions cited above are due to a conceived need to categorise all words into the known word categories of noun, adjective, adverb and secondly, in the case of Stanković, a misconceived notion that a linguistic system, i.e. the grammatical pattern of a language, is unchangeable.

In my view, a more interesting common feature of these E1s (left-hand elements), is their lack of declension which incorporates them into the word category of the E2 (right-hand element), thus making them a part of a compound, where the E2, being declinable, quite unsurprisingly, renders the word category of the first element in the compound irrelevant.

Despite the different ways of describing and explaining the construction N<sub>UNDECL</sub>N I will, as mentioned, draw from *The Oxford Handbook of Compounding* (Lieber and Štekauer 2011) in which it says:

---

<sup>203</sup> Original: „У синтагматском систему српског језика нема модела у коме се једна именица прикључује другој и у непроменљивом облику врши привеску функцију.“

<sup>204</sup> Original: „Riječ je o najčešće dvočlanim tvorbama sastavljenima od imenica koje se uvijek javljaju u određenom, nepromjenjivom redoslijedu i od kojih prva određuje drugu. Prvi je član redovito nepromjenjiv“ (Horvat and Štebih-Golub 2010: 9)



“the most important of the criteria for distinguishing compounds [are]: (i) stress and other phonological means; (ii) syntactic impenetrability, inseparability, and inalterability; and (iii) the behaviour of the complex item with respect to inflection.” (Lieber and Štekauer 2011: 8)

I believe that by applying the criterion of syntactic impenetrability, inseparability, and inalterability we are able to group the [N<sub>UNDECL</sub>N]<sub>N</sub>-constructions with otherwise recognised compounds in Serbian and Croatian.

Whereas criterion (i) and (iii) relate to the same issues as Klajn's criterion 3 and 4 (p.182), criterion (ii), which is inspired by English morphological lexicology (Lieber and Štekauer 2011: 11), makes it possible to distinguish the compound *kontakt grupa* [N<sub>UNDECL</sub>N]<sub>N</sub> from the idiomatic NP *kontaktna grupa* [A<sub>DECL</sub>+N]<sub>NP</sub>. It is, in my view, this syntactic impenetrability, inseparability and inalterability that is one of the key reasons why speakers of Serbian and Croatian so readily use these replicated [N<sub>UNDECL</sub>N]<sub>N</sub> compounds and even form new ones using the same pattern: *nana čaj* (< 'mint tea'), *jagoda cocktail* (< 'strawberry cocktail'). The syntactic impenetrability, inseparability and inalterability may be illustrated by the following three examples of [N<sub>UNDECL</sub>N]<sub>N</sub>-compounds:

[N<sub>UNDECL</sub>N]<sub>N</sub> *kontakt osoba* - 'contact person'  
 [N<sub>UNDECL</sub>N]<sub>N</sub> *čarter let* - 'charter(ed) flight'  
 [N<sub>UNDECL</sub>N]<sub>N</sub> *sladoled kocka* - 'ice cream cube'

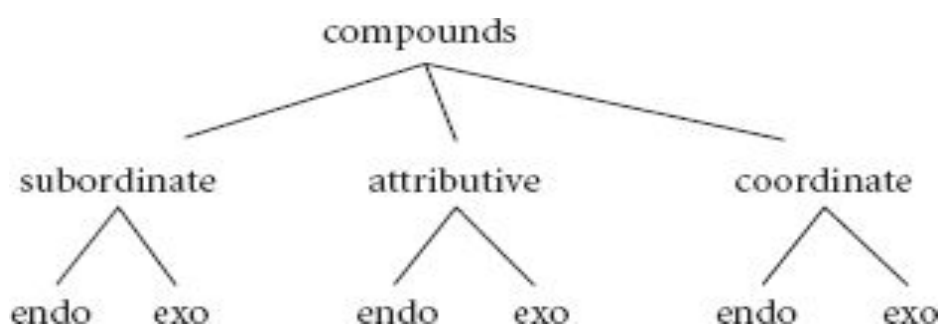
The elements in *kontakt osoba*, *čarter let* and *sladoled kocka* are *inseparable*, i.e. it is not possible to insert a (modifying) element between the head and the non-head of the construction. In the two consulted electronic text corpora (HNK 2.5 and SrpKor 2013) there are no instances of *kontakt X osoba*, *čarter X let* or *sladoled X kocka* (X being an inserted wordform). Furthermore, the first element is *unalterable*, i.e. undeclined, and when in this construction it is even indeclinable regardless of whether it is a replicated wordform (as *čarter*, *kontakt*) or a domestic one (*sladoled*), all declinable nouns in other constructions.

In chapter 5 „A Lexical Semantic Approach to Compounding“ of *The Oxford Handbook of Compounding* (2011), Lieber uses a classification of compounds defined and described by Scalise and Bisetto in chapter 3 „The Classification of Compounds“ where both the syntactic and seman-

tic relationship between the elements and the semantic referential content are included. Their classification of compounds is not language dependent and their goal is that it should be universally valid<sup>205</sup> (Scalise and Bisetto 2011: 34). Syntactically, they apply not only a distinction between coordination and subordination between the elements of the compound; they add a third category – attribution.

Semantically we distinguish between endocentric and exocentric compounds: „Endocentric compounds are, of course, those in which the compound as a whole is a hyponym of its head element. Exocentric compounds are those in which the compound as a whole is not a hyponym of its head.” (Lieber 2011: 89). Thus we end up with six possible compound structures.

Figure 4.2 – Six possible compound structures



(Scalise and Bisetto 2011: 45)

### ***Coordinate compounds***

„Coordinate compounds are formations whose constituents are connected by the conjunction 'and'.” (Scalise and Bisetto 2011: 46).<sup>206</sup>

This indicates that, in a semantic sense, both (or all) elements in a coordinate compound are semantic heads but in order for it to qualify as a compound one of the elements (in Croatian and Serbian, the E2) is syntactically the head whereas the other element (E1) morphosyntactically is the less important element, i.e. it is in form unalterable and therefore does not signal any syntactic role to other elements than the syntactic head of the compound.

<sup>205</sup> This classification has been tested on 23 languages, one of these being Serbo-Croatian (Bisetto and Scalise 2011: 34, footnote 1)

<sup>206</sup> Other scholars have called this type of compound *copulative compounds*.

In Croatian and Serbian, there are few examples resembling the pattern in the English endocentric coordinate nominal compounds such as *poet-doctor*, *boy-actor*.

(1) *pjesnik-pjevač* 'poet-singer'<sup>207</sup>

(1a) *sa pjesnikom-pjevačem*

PREP N.INST

N. INST

'with poet-singer'

'with the singer-songwriter'

As shown in (1a), in Serbian and Croatian both elements in this collocation are declinable, i.e. both elements signal, in congruency, a syntactic role to other elements in the phrase and therefore are not compounds according to the present classification of compounds. They have also been identified as replicated lexical patterns (Horvat and Štebih-Golub 2010: 9) and because the internal order of the elements is unalterable (i.e. *pjesnik-pjevač* exists whereas I have found no indications that *\*pjevač-pjesnik* does) this type of [N N]-constructions are not uninteresting here because they may be examples of an adaptation of the model compound pattern [N<sub>UNDECL</sub>N]<sub>N</sub> found in e.g. *boy-actor* to the [N N]-pattern in *pjesnik-pjevač*.

Whereas the elements in an endocentric type refer to the same denotatum (a man who is both a poet and a singer), the exocentric type refers either to (a) the relation between the denotata or (b) to a third denotatum. Both these types of exocentric coordinate compounds exist in Serbian and Croatian:

(2) *jugoistok* - 'northeast' *Austrougarska*, - 'Austria-Hungary',  
*vukodlak* - 'wolf-hair' (werewolf) (nominal compounds with a third denotatum).

(3) *srpsko-ruski* - 'Serbo-Russian' (adjectival compound conveying a relation between the two denotata).

### **Attributive compounds**

Attributive compounds „consist in a noun-head that can be modified by an adjective, [...], by a noun or a verb.” (Scalise and Bisetto 2011: 45). Even when the non-head [E1] is a noun, and therefore an apposition, “the non-head has an ‘adjectival’ function.” (Ibid: 49) and is “acting as an at-

---

<sup>207</sup> Equivalent to 'singer-songwriter'

tribute”(Ibid: 51). This means that the syntactic structure of an attributive compound is that of head (which is a noun) and attribute which may be a noun, an adjective or a verb.

- (4) *bljedolik* - 'pale person' (paleface)  
ATTR+HEAD  
[A + IFX + N]<sub>N</sub>
  
- (5) *crvenokožac* – 'red-skinned person' (redskin)  
ATTR+HEAD  
[A + IFX + N+ SUFFIX –ac]<sub>N</sub>
  
- (6) *instant kafa* – 'instant coffee'  
ATTR+HEAD  
[A<sub>UNDECLN</sub>]<sub>N</sub>
  
- (7) *gala predstava* – 'gala performance'  
ATTR+HEAD  
[N<sub>UNDECLN</sub>]<sub>N</sub>
  
- (8) *gej zajednica* – 'gay community'  
ATTR+HEAD  
[N/A<sub>UNDECLN</sub>]<sub>N</sub>
  
- (9) *Milford čaj* - 'Milford Tea'<sup>208</sup>  
ATTR+HEAD  
[PN<sub>UNDECLN</sub>]<sub>N</sub>
  
- (10) *Sava Centar* – 'The Sava Centre'<sup>209</sup>  
ATTR+HEAD  
[PN<sub>UNDECLN</sub>]<sub>N</sub>

Endocentric attributive compounds where the non-head ascribes a property to the head are common in as well conventionally recognised compounds (4) and (5) but also in the [N/A<sub>UNDECLN</sub>]<sub>N</sub>-compounds as (6-8).

---

<sup>208</sup> *Milford* is the name of the tea-producing company.

<sup>209</sup> *Sava* is the name of the river, on whose banks this “congress, cultural and business center” is situated, ('Sava Centar' 2016)

Example (9) and (10) stand out by being names of companies, products or businesses. Such compounds are a particularly productive subtype within the endocentric attributive compounds, with an undeclined proper noun (PN) as E1 (Starčević 2006: 650)<sup>210</sup>. However, I have chosen not to query the electronica corpora for compounds consisting of one or two proper nouns because the focus of the investigation is to ascertain the usage norm of the speakers and names of products, businesses, companies, festivals, events, etc. are not representative of this usage norm. The motivation behind their formation may even be just the opposite, i.e. a tendency to go against the norm to fulfil a commercial need to create attention and establish recognition of a specific brand.

The English model for examples (4) 'paleface' and (5) 'redskin' are exocentric attributive compounds (Scalise and Bisetto 2011: 46). Thus, these compounds refer to a denotatum which is not a hyponym of its head noun ('face' or 'skin'). Exocentric attributive compounds are, however, not readily formed in Serbian and Croatian. So, while replicating the lexical pattern of e.g. 'redskin', the head noun – 'skin' (*koža*), is converted so that the head becomes a hypernym: *kožac* ('person with skin') of the compound *crvenokožac* ('person with red skin') In other words, the lexical pattern (RED + SKIN) of an exocentric attributive compounds is replicable into Croatian and Serbian but the exocentricity is avoided by altering the head of the compound.

- (11) *haljina-košulja* – 'dress-shirt' (shirt dress)

HEAD + ATTR

N + N

- (12) *čovjek-žaba* - 'man frog' (frogman)<sup>211</sup>

HEAD + ATTR

N + N

Finally, we have examples of attributive compound-like constructions as those in (11) and (12), which have a striking similarity to coordinate constructions as in (1) *pjesnik-pjevač*. The difference between them, applying Scalise and Bisetto's categorisation, lies in the relationship between

---

<sup>210</sup> Starčević (2006) terms this preposed apposition - *predapozicija*

<sup>211</sup> (10) and (11) are mentioned as examples of loan translations (replicated lexical pattern) in (Horvat and Štebih-Golub 2010: 9)) Another example: *čovjek-mrav* – 'man-ant' (<Ant-man)

the two elements and the denotatum. Whereas *pjesnik* ('poet') and *pjevač* ('singer') refer to one denotatum (the same person), *žaba* ('frog') and *košulja* ('shirt') are descriptions, i.e. ascribe properties to their head elements, meaning that the 'frogman' is a man who *acts* like a frog without being a frog and the 'shirt dress' is a dress which looks like a shirt. Both elements are declined, so these constructions will not be considered compounds in Serbian and Croatian even though they are compounds in the model language – English. The construction HEAD + ATTR is not foreign to Serbian and Croatian, but is generally described as a case of nominal apposition (Barić et al. 1995: 563) or nominal attribution (Piper and Klajn 2013: 306) and are as such not considered compounds nor semi-compounds but NPs. Typical examples in the mentioned grammars are:

- (13) *čovjek očajnik* – 'man desperado' (desperate man)

HEAD + ATTR – a man who is also a desperado

N + N

- (14) *grad heroj* – 'city hero' (heroic city)

HEAD + ATTR – a city which is also a hero

N + N

### ***Subordinate compounds***

„Compounds are to be defined as 'subordinate' when the two components share a head-complement relation.“ (Scalise and Bisetto 2011: 45)

- (15) *pismonoša* – 'letter bearer' (bearer of letters/mailman)

COMPL+HEAD

[N+IFX+ N]<sub>N</sub>

- (16) *internet korisnik* – 'internet user'

COMPL + HEAD

[NUNDECLN]<sub>N</sub>

- (17) *nana čaj* - 'mint tea'

COMPL + HEAD

[NUNDECLN]<sub>N</sub>

- (18) *jagoda cocktail* – 'strawberry cocktail'

COMPL + HEAD

[NUNDECLN]<sub>N</sub>

(15)-(18) are all endocentric compounds.

Not only compounds where the head is a deverbative as *-noša* (< *nos-iti* – 'to bear') and *korisnik* (< *korist-iti* – 'to use') are classified as subordinate compounds, but also compounds where there is a „strong 'of relation'“ as in *mushroom soup* or *apron string*. The non-head is a part of the head, i.e. is a complement to the head, viz. a soup made *of* mushrooms / a soup with a flavour *of* mushrooms, a string *of* an apron/a string resting on an apron/a string threaded into an apron) and not, as in attributive compounds, a property of the head. (Scalise and Bisetto 2011: 45; 50). This analysis is applicable to (17) and (18) where the mint is not a property of the tea, but rather a part or flavour *of* the tea and the strawberry is not a property of the cocktail but a *part of* or an *ingredient* in the cocktail.

Exocentric subordinate compounds are compounds in which the subordinate complement is in E2-position and determines the word category of the compound. Scalise and Bisetto exemplify this subtype with *pickpocket* and *killjoy* as well as with *lavapiatti* (Ital. 'wash' + 'dishes' > 'dishwasher') (Scalise and Bisetto 2011: 45).

Such compounds exist in Serbian and Croatian, but I have not found any evidence of this being a replicated pattern and the word-formational pattern in these compounds is today viewed as unproductive (Klajn 2002: 86). In the Serbian and Croatian linguistic traditions, they are called 'imperative compounds' (*imperativne složenice*) as most of the verbal superordinate heads are recognised as imperative forms in 2.prs. sing. (as could the English *pick* and *kill* in *pickpocket* and *killjoy*).

(19) *secikesa* – 'cutpurse' *seci* – 'cut', *kesa* – 'purse'

HEAD+COMPL

[V + N]<sub>N</sub>

Interestingly, the case of the complement is not governed by the verb's valence, as it would be in a VP as in (20).

- (20)    *Seci*                      *kesu!*  
           VB.IMP.2.PRS.SING   N.ACC.SING  
           ‘Cut [the] purse!’

#### 4.4.2 Contact-induced attributive, subordinate and coordinate compounds in Serbian and Croatian

We have now seen that Serbian and Croatian may boast at least five out of Scalise and Bisetto’s six categories of nominal compounds or nominal compound-like constructions:

- Coordinate endocentric: (1) *pjesnik-pjevač*  
 Coordinate exocentric: (2) *jugoistok, vukodlak*  
 Attributive endocentric: (4) *bljedolik*, (7) *gala predstava*,  
   (12) *čovjek-žaba*  
 Attributive exocentric: -  
 Subordinate endocentric: (15) *pismonoša*, (16) *internet korisnik*,  
   (17) *nana čaj*  
 Subordinate exocentric: (19) *secikesa*

As noted above, the exemplified coordinate endocentric compounds ((1) *pjesnik-pjevač*) do not meet the criterion that the first element be undeclined, though they might very well be contact-induced, i.e. the lexical pattern ‘poet-singer’ could be construed as being replicated (through loan translation or rendering) of the English model *singer-songwriter*. The indeclinability-criterion also leads us to exclude attributive endocentric constructions such as (12) *čovjek-žaba* (‘frogman’). Among the coordinate exocentric compounds ((2) *jugoistok, vukodlak*) which are formed as ‘true’ nominal compounds, i.e. they have only one stress, their two elements are linked with the linking element *-o-*, I have hitherto not found any clearly contact-induced examples, which has led me to exclude all coordinate compounds from further investigation in this study.

The endocentric attributive nominal compounds such as (7) *gala predstava* and the endocentric subordinate nominal compounds such as (16) *internet korisnik*, will presumably be most numerous among the contact-induced compounds, and it shall be interesting to see whether the empirical data will reveal any contact-induced exocentric compounds among the attributive and subordinate kind.

Whether an innovation has been induced by contact or not becomes, as we have seen, increasingly difficult to assess the further we move along



the cline from lexical matter towards grammatical pattern. I will, nonetheless, categorise the contact-induced compounds I find into four categories: *Replicated matter*, *Replicated and inherited matter a replicated lexical pattern*, *Inherited matter in a replicated lexical pattern*, *Replicated word-formational pattern* along this cline, Cline e, where the replication process inflicts increasingly on the linguistic system of the replica language the further we move to towards the right along the cline.

*Cline e*

lexical	>	>	>	>	>	grammatical
content	>	>	>	>	>	function
matter	>	matter + lexical pattern			>	lexical pattern
					>	word-formational pattern

In Table 4.4 below and in the subsequent subsections the four categories are exemplified and further categorised as having either an attributive or a subordinate endocentric internal structure.

Table 4.4 – Replicated matter and pattern in nominal compounds

<b>Replicated matter</b>	
Attributive, endocentric	Subordinate, endocentric
<i>kontakt grupa</i> – ‘contact group’ <i>džez muzika</i> – ‘jazz music’ <i>rok muzika</i> – ‘rock music’ <i>basket parket</i> – ‘basketball floor’ <i>tampon zona</i> – ‘buffer zone’ <i>duhankesa</i> – ‘tobacco pouch’	<i>šok terapija</i> – ‘shock therapy’ <i>horor-šou</i> – ‘horror show’ <i>rum koktel</i> – ‘cocktail containing rum’
<b>Replicated and inherited matter in a replicated lexical pattern</b>	
Attributive, endocentric	Subordinate, endocentric
<i>internet veza</i> – ‘Internet connection’ <i>parking prostor</i> – ‘parking lot’	<i>internet korisnik</i> – ‘Internet user’ <i>pop pevač</i> – ‘pop singer’
<b>Inherited matter in a replicated lexical pattern</b>	
Attributive, endocentric	Subordinate, endocentric
<i>kolodvor</i> – ‘railway station’ (< Ger. Bahnhof) <i>poklon paket</i> – ‘gift package’	<i>nana čaj</i> – ‘mint tea’
<b>Replicated word-formational pattern</b>	
Attributive, endocentric	Subordinate, endocentric
<i>eskimo-sladoled</i> – ‘ice cream pop-sicle’ <sup>212</sup> <i>sladoled kocke</i> – ‘ice cream cubes’ <i>bobi-štapíci</i> – ‘grissini’ <sup>213</sup>	

### **Replicated matter**

Contact-induced compounds, which consist entirely of replicated lexical matter, adapted phonologically, morphologically and possibly orthographically correspond to the kind of linguistic transfer, which Heine and Kuteva, quoted in subsection 3.3.1 above, label *Form-meaning units or combinations of form meaning units*. Because contact-induced, replicated compounds may have been replicated entirely in form with or without

<sup>212</sup> Explained by Ćirilov (1982:61) as “sladoled na drvenim štapícima” (ice cream on small wooden sticks)

<sup>213</sup> Ćirilov (1982: 27)

orthographic adaptation as in the replication of ‘comeback’, which in the empirical data is orthographically rendered in the following six different ways: *come-back*, *come back*, *comeback*, *kam-bek*, *kambek* and *kam bek*, they may also consist of previously replicated matter (adopted matter) which, induced by the combination in a model language, is also combined in the replica language. In any case, on *Cline e* (above) this kind of replication should be placed leftmost as it concerns only lexical matter.

Arguably, the ‘comeback’-example of replicated matter should not be listed under the heading of compounds at all, because the composition of the two elements *come* and *back* did not occur in the replica language, which means that it must have been replicated as a simplex lexeme.<sup>214</sup> It and other compounds were formed as compounds in the model language. Notwithstanding, I include them here because the fact that speakers of Croatian and Serbian (re)analyse them as consisting of two (lexical) elements, supposedly because they recognise the individual elements in other replicated matter, may lead them to form new compounds or replicate more compounds with recognisable elements as for instance *bek/back* (‘back’) in replicas of ‘feedback’, ‘flashback’ and ‘playback’.

Haspelmath (2009) explains this phenomenon in a similar way stating: “However, when a language borrows multiple complex words from another language, the elements may recur with a similar meaning, so that the morphological structure may be reconstituted.” (Haspelmath 2009: 37). The first element in this sort of compound is, however, rarely (re)analysed as a full-fledged adjective or noun (a lexeme), but rather as a kind of prefix or “semiword”, which I will return to in “4.5 Semiwords and affixes”.

A compound consisting of replicated matter may also be the result of word-formation in the replica-language itself. This is the case when two lexemes have been replicated, and then combined in a compound by speakers of the replica-language. Both elements are replicas and have entered into the replica language at an earlier stage as replicated individual lexemes, which only later have been combined in a compound as e.g. *basket parket* (‘basketball floor’) where both *basket* and *parket* are replicated from other languages<sup>215</sup>. The combination of these two specific content items has, in other words, happened language-internally. The

---

<sup>214</sup> As in the replica of the German compound for ‘sandpaper’: *šmirgl-papir* < Ger. Schmirgelpapier

<sup>215</sup> *basket* < (Eng.) basket(ball), *parket* < (Fr.) parquet

way they are combined, i.e. [N<sub>UNDECLN</sub>]<sub>N</sub> is in fact, as we have seen, induced by contact but it is then only an instance of pattern replication. However, on the surface, i.e. the individual pieces of matter, are also replicas and for this, practical, reason they are grouped with compounds which have been replicated as compounds, as e.g. *kontakt grupa* (< ‘contact group’). If it cannot be established that the compound is a replica we must assume that the compounds as e.g. *basket parket* have been formed in the replica language, albeit with formerly replicated lexical matter. Haspelmath (2009: 39) terms this type of word-formation “loan-based creation”, i.e. words formed in language L1 with matter that was previously replicated from any language (Lx) other than L1. So, the category in Table 4.4 labelled “Replicated matter” we may find compounds which also belong in the category labelled “Replicated word-formational pattern”, i.e. only the pattern has been replicated when forming the compound, and, in these overlapping instances, the individual elements have been replicated at an earlier date.

The semantic and syntactic features of the model compounds, whose replicas are not compounds (e.g. *kam-bek*) may be analysed within the trichotomy of coordinative, attributive and subordinate compounds outlined and explained above, but it would not serve our purpose as they do not, in their replicated form, constitute compounding in Serbian and Croatian and therefore do not inform us about the possible semantic and syntactic/word-formational constraints on compounds formed in Serbian and Croatian. In turn, this means that at this stage they do not represent an innovation in the word-formational possibilities of Serbian and Croatian. They do, however, constitute a model on which compounds may be formed in Serbian and Croatian which is what we witness in *basket parket*, *nana čaj* and *sladoled kocke*.

The semantic and syntactic features of the loan-based creations (*basket parket*) and replicated compounds where both elements also exist as independent content items in the replica language (*šok terapija*) are intriguing as they, along with compounds formed entirely out of inherited<sup>216</sup> lexical matter (*nana čaj*) or out of a combination of inherited and replicated lexical matter (*internet veza*), may reveal which restrictions on the formation of compounds (of domestic and adopted matter), exist

---

<sup>216</sup> “inherited matter” is a term that covers both the domestic and the adopted matter, i.e. originally replicated matter which is considered to be so assimilated into the RL, that it ‘passes’ as domestic, for instance *čaj* (tea), *duhan* (tobacco) or *kava/kafa* (coffee)

in the Serbian and Croatian usage norm and whether there are differences the two norms exhibit any differences in this respect.

***Replicated and inherited matter in a replicated lexical pattern***

Compounds formed out of a combination of inherited and replicated matter in accordance with a replicated lexical pattern belong to a combination of two kinds of linguistic transfer, outlined by Heine and Kuteva (2005: 2), that is, a combination of *Form-meaning units or combinations of form meaning units* and *Syntactic relations, that is, the order of meaningful elements*. On *Cline e* (above), we have now moved one step away from the pure matter replication and are now dealing with word-formations, constructions where the lexical pattern from the model language appears in compounds which have been formed in the replica language. In other words, it is no longer a question of adapting foreign matter (phonologically and otherwise), but rather an innovative rendering of concepts hitherto unexpressed in this manner by inherited means. It is, indeed, an instance of partial loan creation, or a combination of *Loan Creation* and *Loan Translation*, introduced in Figure 4.1, p. 164.

These compounds are also called semi-calques and hybrids (Štebih 2008: 248) which refers to the fact that only half of the compound has been translated (gone through calquing) and that the other half is replicated matter. Let us recall that the term *hybrid*, in modernity has a distinct negative connotation connecting it to impurity, disorder and interference as opposed to the orderly original, untainted, pure, authentic and domestic lexeme, cf. p. 72. In order to avoid these negative connotations, I will, if it can be established that the compound is a replication of an equivalent compound in a model language, use the term semi-calque to refer to compounds which are instances of (the quite long description) “replication of matter combined with translation of lexical matter within a replicated lexical pattern”. If it cannot be established that the compound or compound-like construction is a replica, we must assume that the compound has been formed in the replica language and view it as a semi-loan-based creation, which differs from the loan-based creations mentioned in the previous section only in that they consist of both inherited and replicated matter.

It may be very difficult to determine whether a compound has been replicated in its entirety, i.e. whether it was (b) formed in the model language or it has (a) been formed within the replica language:

21) *grupi-devojka* – ‘groupie-girl’

a) *grupi* ‘groupy’ was presumably, because of its phonetics, replicated from English and subsequently this undeclined replicated content item was combined with the domestic *devojka* ‘girl’: (groupy > *grupi*) + *devojka* > *grupi-devojka*

b) *grupi-devojka* is a semi-calque (a combination of Loan rendering and Loan translation) as the lexical pattern has been replicated along with half of its lexical matter. The other half has been translated: groupie girl > *grupi-devojka*

Whether analysis (a) or (b) is the more valid does not change the fact that the word-formational pattern [N<sub>UNDECLN</sub>]<sub>N</sub> is present in Serbian and Croatian. We may conclude that the pattern must be replicated because the word-formational pattern [N<sub>UNDECLN</sub>]<sub>N</sub> is not considered to be a Croatian and Serbian word-formational pattern by leading, canonical linguists in both Croatia and Serbia such as Stevanović (1964), Maretić (1924) and Babić (1971). These linguists have even, for precisely this reason, advised against the use of such constructions (Barić 1980: 35–38).

The fact that some of the compounds are not replications of equivalents in a model language does not render them irrelevant in contact linguistics, quite the opposite. In such cases it is clearly a case of replication only of the word-formational/syntactic pattern. In other words the word-formational pattern of languages of contact (such as Turkish, German and English) has become a word-formational pattern in Serbian and Croatian, which we will return to shortly (p.201).

Additionally, I will in my investigation, group the compounds according to whether the E1 is replicated or not because another aspect of these compounds is the indeclinability of the E1.

In Table 4.4 above, I categorised a number of endocentric compounds according to the role of the non-head (attribute, subordinate). I initially attempted also to categorise the compounds according to their formational type, (replicated compound, loan-based, semicalque or semi-loan-based creation) but it turned out to be extremely difficult, not to say impossible, because we may, without any chance of getting a proper answer, ask the question whether a particular compound, e.g. *rok*

*muzika* is a replicated compound or a loan-based creation of ‘rock music’ or whether *pop pevač* is a semicalque of *pop singer* or a semi-loan-based creation. Our only solid evidence of loan-based creation is the non-existence of a model in the supposed model language, as we see it in *basket parket*. So, the compounds in Table 4.4 are only grouped as either consisting of i) only *replicated matter*, ii) a blend of *replicated and inherited matter in a replicated lexical pattern*, iii) only *inherited matter in a replicated lexical pattern* or iv) *inherited matter in a replicated wordformational pattern*. The grouping of some of the compounds in Table 4.4 above, provokes the question: When does a replicated E1 cease to be ‘foreign’ and become ‘domestic’. Or: When does any replicated matter cease to be merely *adapted* and become *adopted*? I believe that many a speaker of Serbian or Croatian would treat and regard both *kontakt* and *grupa* in *kontakt grupa* as a ‘domestic’ or ‘inherited’ lexemes and hence categorise *kontakt grupa* as a compound consisting entirely of non-replicated matter. In *duhankesa* (‘tobacco pouch’) both *duhan* (< Tur. ‘dūhan’ < Arab. ‘duhan’) and *kesa* (< Tur. ‘kese’) ‘feel’ domestic and many other E2s, as e.g. *zona* in *tampon zona* also ‘feel’ domestic rather than ‘foreign’. The categorisation is problematic because it is notoriously difficult to base analyses on “das Sprachgefühl” since this ‘feeling’ may vary from person to person.

It may also prove difficult to subcategorise not only these but all the compounds according to the role of the non-head (attributive and subordinate). With every compound we have to ask ourselves the question: Is the non-head a property that merely describes the head or is it rather a complement of the head. In other words; is *horror* an attribute to *show* in *horror show* (cf. p. 195) as is *funny* in ‘a funny show’ or is *horror* a complement of its head *show*, as in ‘the show creates horror’? Other compounds are easier to categorise: *šok terapija* (‘shock therapy’) is a therapy which consist in giving electrical shocks, so ‘shock’ is the complement of ‘therapy’, i.e. *šok terapija* is an endocentric subordinate compound.

As useful as these categorisations may be in showing the many different syntactico-semantic relations between the elements in a compound, they prove very difficult to apply in an analysis of all compounds.

### ***Inherited matter in a replicated lexical pattern***

Compounds formed in accordance with a replicated lexical pattern without any replicated matter would belong to a combination of the two kinds of linguistic transfer as defined by Heine and Kuteva (2005) that is; a

combination of “Meanings or combinations of meanings” and “Syntactic relations, that is, the order of meaningful elements.” (Heine and Kuteva 2005: 2). We have now moved one step further up *Cline e* (p. 194) where we find the compounds which constitute replicas of lexical patterns in model languages:

22) *djelokrug* (< Ger. Wirkungskreis)

23) *kolodvor* (< Ger. Bahnhof)

24) *nana čaj* (< ‘mint tea’)

25) *poklon paket* (< ‘gift package’)

These compounds are the result of lexical pattern replication, and are called *loan translations* or *calques*, i.e. a more or less direct translation of a compound in a model language which therefore has no overt traces of the model language but nevertheless the patterning (the order and the syntactic relations) of the involved content items is replicated.

In Serbian and Croatian the increase in such replicated compounds is believed to have begun in the mid 19<sup>th</sup> century and for the early period the model languages are Latin, German and Turkish (Klajn 2002: 38–39). The most dominant of the model languages was German, especially in Croatian, but also in Serbian (Zett 1970: 16).

In this group of calques I have also included compounds such as (24) *nana čaj* and (25) *poklon paket* in which the second element (E2) is a previously replicated (adopted) content item because the E2 is declinable and therefore its etymology is less relevant in the analysis of the investigated innovation in compound formation in Serbian and Croatian.

It is also among the loan translations that we find the replicas of coordinate such as (1) *pjesnik-pjevač* and attributive endocentric compounds such as (17) *čovjek-žaba* which, though being compounds in the presumed model language (‘singer-songwriter’ and ‘frogman’) do not meet the criterion of indeclinability of the first element in the replication and therefore are not considered nominal compounds in Serbian and Croatian.

### ***Replicated word-formatational pattern***

Compounds formed in accordance with a replicated word-formatational pattern without any replicated matter would belong to the kind of linguistic transfer, which Heine and Kuteva (2005: 2) define as: “Syntactic relations, that is, the order of meaningful elements.”

So, furthest up the *Cline e* from lexical towards grammatical innovations within compounds in Serbian and Croatian we find compounds where



only the *word-formatational pattern* has been replicated. In these compounds neither the lexical matter nor the lexical pattern of the compound, i.e. the *content* have been inspired by a model in another language. Only the way in which content items are combined into compounds can be presumed to have been replicated. The increased number and usage of compounds formed according to a replicated lexical pattern, with or without replicated matter, outlined in the previous subsections (e.g. *nana čaj*, *kontakt grupa*, *internet veza*) may have resulted in an increase in compounds formed without any foreign lexical model. In other words, in these instances the number of ways in which the language users of Serbian and Croatian combine two nouns has increased so that new nominal compounds may be formed.

Here, the originally foreign compounding pattern, that is, a compounding pattern found in a model language, becomes an integrated component part of the replica language system, i.e. the structural build-up of Serbian and Croatian.

26) *eskimo-sladoled* - 'ice cream pop-sickles'

27) *bobi-štapíci* - 'grissini'

28) *sladoled kocke* - 'ice cream cubes'

Because [N<sub>UNDECL</sub>N]<sub>N</sub>-compounds have existed in Serbian and Croatian for centuries, though to a much lesser extent, e.g. : *spomen ploča* ('memorial platter'), *paradajz čorba* ('tomato soup') (Stanković 1997: 94–95), the increase of this construction should be viewed as an instance of a contact-induced innovation which Heine and Kuteva (2005) term a "Rise from a minor to a major use pattern" which involves that an existing pattern is used more frequently, in new contexts and may become associated with a new grammatical function (Heine and Kuteva 2005: 45–46)<sup>217</sup>. Šević, a Serb linguist claims exactly this when she says that 'the English model has served as a catalyst of the intense use of this latently present construction in the Serbian language.' (Šević 1996: 85)<sup>218</sup>

Arguably, loan-based creations (*basket parket*) and semi-loan-based creations (*grupi-devojka*) also belong to this category as they have been formed in the replica language. Nevertheless, they constitute a spe-

---

<sup>217</sup> cf. also p. 83

<sup>218</sup> Original: "...engleski model [je] poslužio kao pokretač intenzivne upotrebe ove već latentno prisutne konstrukcije u srpskom jeziku."

cific sub-category because it might be stipulated that they are more easily formed according to this replicated pattern *because* they consist of replicated matter rather than of inherited matter.

### ***Summing up the nominal compounds***

Before proceeding to the empirical evidence of such compounds in the two electronic corpora, I will briefly sum up the categories of contact-induced [N<sub>UNDECL</sub>N]<sub>N</sub>-compounds in Serbian and Croatian.

i) Matter replication where we see compounds replicated in their entirety with (*kontakt grupa*) or without (*kam-bek*) discernable elements in the replica or even a so-called loan-based creation where both elements have been replicated individually (*basket parket*).

(ii) Replication of only the lexical pattern (*nana čaj, djelokrug, kolodvor*).

(iii) The compounds in question may also be a replication of a lexical pattern in which the E1 is also replicated matter, which constitutes a combination of category (i) and (ii) (*čarter let, internet-stranice*).

(iv) Finally, they may belong to the compounds formed out of inherited content items without being a translation of a foreign compound, thus presumably constituting instances of word-formational pattern replication (*sladoled kocke*) or simply a contact-induced increase in usage of an existent grammatical (word-formational) use pattern. I assume that this particular form of pattern replication would not have happened without the extensive matter and pattern replication as in in the category (i) and (ii), which is why I propose to call it *relayed pattern replication*, i.e. a pattern which is replicated not necessarily due to direct contact with another language, the model language, but rather due to the indirect contact via the imported compounds, i.e. compounds which consist of replicated matter as well as replicated pattern.

### ***Normative attitudes towards [N<sub>UNDECL</sub>N]<sub>N</sub>-compounds***

This type of nominal compounds exists in Serbian and Croatian language usage, as revealed by the findings in corpus searches, despite protests and proscriptions from the linguistic authorities in both Croatia and Serbia. Let us recall how the “Institute for Croatian Language and Linguistics” (IHJJ) proscribed *internet-stranice*, moderately recommended the less accepted *internetske stranice* and highly recommended the more acceptable *mrežne stranice* (IHJJ 2015a). In a similar way *The Board of Standardisation of the Serbian language* proscribes *kontakt grupa* and *kompakt disk* and prescribes the adjectivisation of E1 in the replicated compounds *kontakt grupa, kompakt disk* and *marketing miks* forming the

[A+N]<sub>NPS</sub> *kontaktna grupa*, *kompaktni disk* and *marketinški splet* ('contact group', 'compact disc' and 'marketing mix') (Brborić et al 2006: 54; 102). The Croatian *Council for Standard Croatian Language Norm* is harsher in its condemnation of this foreign (English) influence as it objects to the existence of attributive [N+N]<sub>N</sub> – compounds and says that 'by placing one noun in front of another the second noun cannot be determined or modified, viz. that an attributive relation between the two [nouns] may not be expressed in this manner' (VNHSJ 2013: 81)<sup>219</sup>. Immediately after this passage, it is recognised that such objectable constructions do exist in Croatian, that they are a result of English influence and when written they must be written hyphenated, thus marking them as foreign (e.g. *džez-orkestar*), in order to protect the Standard Croatian language system from disturbing implications which would follow from writing them in two words (*džez orkestar*) or unhyphenated (*džez-orkestar*). (Ibid:)

Incidentally, in the written rendering of all four types of compounds (i-iv) there is an orthographic variety, both due to differing orthographic rules depending on script, country and orthographic manuals (orthographically adapted or not, e.g. *чарџер леџ*, *čarter let*, *charter let* ('charter(ed) flight') but also due to the non-adherence to prescribed standardised rules concerning hyphenation, e.g. *internet stranice*, *internet-stranice* ('internet pages'), *vešmašina*, *veš-mašina*, *veš mašina* ('washing machine').

#### 4.4.3 Empirical data

Inspired primarily by examples of [N<sub>UNDECL</sub>N]<sub>N</sub>-compounds provided by both Serbian and Croatian scholars (cited above) I have selected a sample of 18 nouns, which have been pointed out as potential E1s and/or E2s in [N<sub>UNDECL</sub>N]<sub>N</sub>-compounds in the literature. They represent nouns replicated from different model languages and could, according to Croatian normativists, (cf. subsection 4.2.1) be categorised as *partial adaptees* ('tuđice') and *adaptees* ('prilagođenice') and some which I would categorise as *adoptees* (i.e. words of foreign origin which have been completely adopted into Serbian and/or Croatian and which most native speakers do not consider to be foreign) and finally among them we also find a few domestic nouns, i.e. not replicated lexical matter. The two latter catego-

---

<sup>219</sup> Original: "...da se stavljanjem jedne imenice ispred druge ta druga ne može pobliže određivati, da se među njima tako ne može izražavati atributni odnos."

ries (adoptees and domestic) I will continue to subsume under the category *inherited* as opposed to the two previous categories which I subsume under the category of *replicated*.

I have searched the Serbian and the Croatian corpus by querying the following E1s and E2s:

<u>E1</u>	<u>E2</u>
auto	biznis/business
biznis/business	čaj
internet	koktel/cocktail
kamp	muzika/glazba
kontakt	paket
menadžment/management	parket
parking	sladoled
sladoled	šou/show
šoping/shopping	zona

On the following pages individual types of compounds will be presented to exemplify the results and the methodology of the searches and categorisation. All types of compounds found from these queries are listed in appendix A.

#### *Search criteria*

All queries conducted in both corpora were compiled in concordances which I inspected in order to eliminate examples of faulty tagging as well as instances of N<sub>NOM</sub> + N which are not [N<sub>UNDECL</sub>N]<sub>N</sub>-compounds, according to the following criteria:

- The two elements must be nouns.
- The two elements must be part of a compound, defined by applying the criteria of syntactic impenetrability, inseparability, and inalterability, cf. above.
- The E1 must be indeclinable (i.e. in an undeclined, unalterable form, which is identical to the form in nominative, singular).
- The two elements must be nouns in their entire form, i.e. clipped forms such as *avio* in *avio-slужba* ('flight service') are omitted and considered as semiwords.<sup>220</sup>

---

<sup>220</sup> The token *auto* meaning 'car (automobile)' will be considered as an entire and domestic noun, despite its origin as a clipping of the replicated *automobil*, as it is today

- The E1s and E2s must be nouns in Croatian or Serbian, i.e. *mol* in *šoping mol* (shopping mall) or *tok* in *tok šou* (talk show) will be omitted because *mol* and *tok* are not used as nouns in either language.
- The nominal compound must be a common noun, i.e. names of products, businesses, companies, festivals, events, etc. such as (9) *Milford čaj* and (10) *Sava Centar* are omitted.

In samples where it was not straightforward to determine whether a specific E1 or E2 is also used as an independent noun in Croatian or Serbian usage, I conducted additional queries of the specific word-form in one or both corpora. For instance, in SrpKor I found the E2 *šou* combined with the E1 *blokbaster* ('blockbuster'), viz. *blokbaster šou* ('blockbuster show'). In order to determine whether *blokbaster* is indeed an independent noun in the Serbian usage norm, I then queried SrpKor for the noun *blokbaster* and found that *blokbaster* is indeed used not only as an E1 in other compounds (e.g. *blokbaster film*, *blokbaster ideologija*) but also as an independent noun, as is evident the following example, where *blokbaster* is declined as a noun:

*u animiranom blokbasteru "Šrek" (SrpKor)*  
 N MASC.LOC.SING  
 ('in the animated blockbuster "Shrek")

The E2 *fešn* in *fešn šou* ('fashion show') was another example of a 'questionable' noun but *fešn* did not pass the test, as *fešn* did not appear in the corpus as an independent noun. Accordingly, *blokbaster* is registered as one of the 16 different nouns that combine as E1 with *šou* (as E2) in SrpKor, while *fešn* is not<sup>221</sup>.

Compounds which have, with high probability been replicated in their entirety as e.g. *šou biznis* (show business) and therefore not compounds in the replica languages were not omitted as both nouns: *šou* and *biznis* exist as individual nouns in both languages.

Compounds in which one of the elements is a conventionalised acronym, replicated or not, will also be considered [NUNDECLN]<sub>N</sub>-compounds because acronyms as for instance *DNK* in *DNK-paket* ('DNA

---

considered to be an entire noun. However, when the token *auto* has the meaning 'self' it is considered as a semiword.

<sup>221</sup> The element *fešn* could, though, be analysed as a semiword because it holds lexical content without having morphological independence, an issue which is discussed to in section 4.5 below.

package') are treated as nouns and are, when used independently declined and regarded as nouns, as mentioned above (p. 167).

Finally, I did not register the compounds which are names of products, businesses, companies, festivals, events etc. Names of products, businesses, companies, festivals, events, etc. are a specific subtype within word-formation which do not represent the usage norm of the speakers, but rather are formed with respect to other priorities such as a need to be unusual and an aspiration to unique (i.e. against the norm) in order to create attention and establish recognition of a specific brand. This means that occurrences such as *Srbija Sladoled* (name of ice-cream manufacturer) *Amer čaj* (a brand of tea - AMER), *Windows paket* ('Windows package' a software product), *Oprah show* (The Oprah Winfrey Show) are not included in the following analyses.

My analyses of the concordances led me to eliminate a large number of the query results. For instance, of the 175 results including *sladoled* as E2 in SrpKor, not one lived up to the criteria necessary to be categorised as a [NUNDECLN]<sub>N</sub>-compound, which is not also a proper noun.

In the ensuing tables of selected results, the found types of E1 and E2 have been defined as replicated or domestic nouns in order to see how frequently [NUNDECLN]<sub>N</sub>-compounds comprise of only replicated matter, only inherited matter, or both. Domestic nouns (inherited lexical matter) do not only comprise Slavic nouns but also nouns which are, in the words of the Croatian language adviser (HJS), quoted above, by most native speakers not considered foreign at all. It is according to this view that I categorise nouns such as *čaj* and *muzika*<sup>222</sup> as domestic nouns, despite their foreign origin. I will thus also indicate which of the four different categories, presented above 1) *Replicated matter*, 2) *Replicated and inherited matter in a replicated lexical pattern*, 3) *Inherited matter in a replicated lexical pattern* and 4) *Replicated word-formational pattern* the findings represent. As argued, the fourth category represents the final entrenchment of the replicated pattern in the language system of Serbian and Croatian.

---

<sup>222</sup> The noun *muzika* is, however, considered foreign in Croatian, which is why I queried both *muzika* and *glazba* in both corpora.

In order to determine the model language of the replicated nouns, I have consulted various dictionaries, e.g. *Rječnik stranih riječi* ('Dictionary of foreign words') (Klaić 1982).

*A note on orthographic (ir)regularities*

Whether the orthographic rules for writing this type of compound, called 'semi-compounds' (*polusloženice*) in the Croatian and Serbian orthographic manuals and grammars, are followed or not does not tell us anything about the actualisation and implementation of [N<sub>UNDECLN</sub>]<sub>N</sub>-compounds in the usage norms of Serbian and Croatian. The spelling of them can however serve as an indicator of the adherence of these rules, stipulated by the normative works in these language communities.

Though the normative rules of both Serbian and Croatian stipulate that compounds should be written either as one word or hyphenated, it is very clear from the searches that these rules are not followed in the usage norm of the Publicistic functional style in the languages. The compounds with *kamp* as E1 may serve as an illustrative example. One of the most frequent compounds with *kamp* as E1 in both corpora is *kamp kućica* ('weekend cottage'). In SrpKor *kamp kućica* (with no hyphen) appears 16 times and *kamp-kućica* (hyphenated) 14 times, and in HNK *kamp kućica* appears 17 times and *kamp-kućica* 34 times. The inconsistent spelling is evident throughout the empirical data and I interpret the incorrect spelling (i.e. in two words) as a consequence of a number of reasons: 1) Most of the compounds were not considered part of the standard languages when the language users received training in spelling (in school), so it is an example of failed acquisition planning. 2) The language users experience the E1-nouns as independent nouns and therefore put a space between the E1 and the E2. 3) English spelling rules for compounds influence the Croatian and Serbian language users.

Let us now take a look at the overall numbers of different E2-nouns combining with the selected E1-nouns in the two corpora.

Table 4.4.1 - The nouns queried as element 1 in  $[N_{UNDECL}N]_N$ -compounds.

E1-noun	no. of different E2s		Translation
	SrpKor	HNK	
internet	384	184	<i>Internet</i>
auto	149	134	<i>car</i>
biznis/business	108	23	<i>business</i>
šoping/shopping	30	18	<i>shopping</i>
kontakt	27	25	<i>contact</i>
parking	24	16	<i>parking-lot</i>
kamp	13	10	<i>camping</i>
sladoled	4	1	<i>ice-cream</i>
menadžment/management	1	4	<i>management</i>

Excluding *sladoled* and *menadžment/management* as E1-nouns, since I found too little evidence in either corpus to be able to conclude anything but their very scarce use, the generalisation is that in SrpKor we observe a wider use of  $[N_{UNDECL}N]_N$  than in HNK. This is most evident in  $[N_{UNDECL}N]_N$ -compounds with *internet* and *biznis/business* as E1. This result indicates a stronger tendency towards using such compounds in Serbian than in Croatian.

Whether the listed E1-nouns combine more frequently with replicated or inherited E2-nouns, seems to be dependent on the specific E1-noun. And, the percentage distribution of replicated and domestic E2-nouns with the individual E1-nouns (indicated in Table 4.4.6, below) do not seem to constitute a difference between the Serbian and Croatian usage norm.

For instance, *shopping/šoping* as E1 is in SrpKor found in compounds with 30 different E2s of which 11 are domestic nouns, 3 whose model language (ML) is English (E), 5 are registered as coming from French (F), 1 from Italian (I) and 10 which has Classical Greek (H) or Latin (L) as the original model language. All the replicated E2s also exist in English which may, as indicated in section 4.3, work as a catalyst, that increases the usage of the replicated matter.



Table 4.4.2 - Compounds with šoping as E1 and a domestic E2 in SrpKor

<b>Domestic E2s</b>	<i>Translation</i>
šoping groznica	<i>shopping fever</i>
šoping izlog	<i>shopping window</i>
šoping pohod	<i>shopping attack (campaign)</i>
šoping ulica	<i>shopping street</i>
šoping voz	<i>shopping train</i>
šoping hram	<i>shopping temple</i>
šoping navika	<i>shopping habit</i>
šoping grad	<i>shopping city</i>
šoping torba	<i>shopping bag</i>
šoping zabava	<i>shopping fun</i>
šoping put	<i>shopping travel</i>

Table 4.4.3 - Compounds with šoping as E1 and a replicated E2 in SrpKor

<b>Replicated E2</b>	<b>ML</b>	<i>Translation</i>
šoping servis	E	<i>shopping service</i>
šoping spot	E	<i>shopping video</i>
šoping vikend	E	<i>shopping weekend</i>
šoping amater	F	<i>shopping amateur</i>
šoping tura	F	<i>shopping tour</i>
šoping hotel	F	<i>shopping hotel</i>
šoping turista	F	<i>shopping tourist</i>
šoping sezona	F	<i>shopping season</i>
šoping maraton	H	<i>shopping marathon</i>
šoping program	H	<i>shopping programme</i>
šoping zona	H	<i>shopping zone</i>
šoping lista	I	<i>shopping list</i>
šoping agencija	L	<i>shopping agency</i>
šoping asistenat	L	<i>shopping assistant</i>
šoping centar	L	<i>shopping centre</i>
šoping festival	L	<i>shopping festival</i>
šoping globalizacija	L	<i>shopping globalisation</i>
šoping kultura	L	<i>shopping culture</i>
šoping destinacija	L	<i>shopping destination</i>

All the compounds are endocentric as expected<sup>223</sup>, and all of the compounds, save one: *šoping globalizacija* ('shopping globalisation', i.e. the globalisation of shopping) are attributive compounds.

attributive endocentric

*šoping hram*

ATTR+HEAD

[NUNDECLN]<sub>N</sub>

'shopping temple'

subordinate endocentric

*šoping globalizacija*

COMPL+HEAD

[NUNDECLN]<sub>N</sub>

'shopping globalisation'

When comparing with the sample found in HNK (Table 4.4.4 and 4.4.5 below), we find that the compounds, comprising 19 nouns (E2s) with which *shopping/šoping* (E1) combines, have a different spread of domestic (11) and replicated (7) E2s, and only comprise attributive, endocentric compounds.

---

<sup>223</sup> cf. Table 4.4, p.195

Table 4.4.4 - Compounds with shopping as E1 and a domestic E2 in HNK

Domestic E2	Translation
shopping izlet	<i>shopping excursion</i>
shopping izletnik	<i>shopping excursionist</i>
shopping lanac	<i>shopping chain</i>
shopping putovanje	<i>shopping travel</i>
shopping središte	<i>shopping centre</i>
shopping usluga	<i>shopping service</i>
shopping groznica	<i>shopping fever</i>
shopping ponuda	<i>shopping offer</i>
shopping ludilo	<i>shopping frenzy/madness</i>
shopping stranica	<i>shopping page</i>
shopping stanka	<i>shopping break (pause)</i>

Table 4.4.5 - Compounds with shopping as E1 and a replicated E2 in HNK

Replicated E2	ML	Translation
šoping meka	A	<i>shopping mecca</i>
šoping turist	F	<i>shopping tourist</i>
šoping tura	F	<i>shopping tour</i>
šoping turizam	F	<i>shopping tourism</i>
šoping špica	G	<i>shopping rush-hour</i>
šoping lista	I	<i>shopping list</i>
šoping centar	L	<i>shopping centre</i>

The difference in the distribution of domestic and replicated E2-nouns could be viewed as yet another piece of evidence of the divergent tendencies of the Croatian and Serbian usage norms concerning actualisation and implementation of inherited or replicated lexical matter, as is the case with the usage norm concerning replicated nouns and their domestic equivalents (substitutions). However, the compounds with *parking* as E1-noun show the opposite tendency, namely a Croatian tendency to combine with replicated (10) rather than with inherited lexical matter (6) and a Serbian tendency to combine with inherited (14) rather than replicated lexical matter (10). Therefore, I would rather stipulate that the results in Table 4.4.6 show that, when a E1-noun (*internet*, *auto*, *shopping*, *business*, *kontakt*, *parking*, *kamp*) combines with other nouns in [NUNDECLN]<sub>N</sub>-compounds there is no clear difference between the Serbian and Croatian usage norms. In 5 out of 7 (*internet*, *auto*, *business*, *kontakt* and *kamp*) the percentages even indicate that these E1s in SrpKor and HNK combine with a similar percentage of replicated and domestic nouns, i.e. *internet* combines with approximately

as many domestic as replicated nouns, *auto* combines with 1/3 replicated and 2/3 domestic nouns, *business* and *kontakt* combine with approximately 2/3 replicated and 1/3 domestic nouns and finally *kamp* combines more often with domestic than with replicated nouns (in 10 out of 13 (SrpKor) and 9 out of 10 instances (HNK)). Even though the total number of different E2s is quite small with regards to *kamp* (13 and 10) and not very large with regards to *kontakt* either (27 and 25) they do show us that the queried E1s combine with approximately the same number of different E2s in both corpora and they also show us that the distribution of replicated and domestic is similar, cf. Table 4.4.6, below.

Table 4.4.6 – Distribution of replicated and domestic E2s

Replicated (R)/Domestic (D) nominal E2s					
Queried E1-nouns	SrpKor		HNK		Translation
	R/D	R/D	R/D	R/D	
	no. of types	%	no. of types	%	
internet	203/181	53/47	89/96	48/52	<i>internet</i>
auto <sup>224</sup>	40/109	27/73	44/90	33/67	<i>car</i>
biznis/business	73/35	68/32	15/08	65/35	<i>business</i>
šoping/shopping	19/11	63/37	07/11	39/61	<i>shopping</i>
kontakt	18/09	67/33	15/10	60/40	<i>contact</i>
parking	10/14	42/58	10/06	62/38	<i>parking-lot</i>
kamp	03/10	23/77	01/09	10/90	<i>camping</i>

Among the many [N<sub>UNDECLN</sub>]<sub>N</sub>-compounds with *internet* as E1-noun I have, as expected, not found any exocentric compounds, but among the endocentric compounds we find both attributive compounds, where *internet* modifies the E2 e.g. *internet oblak* ('internet cloud'), *internet veza* (internet connection), *internet sustav* (internet system) and subordinate compounds, in which *internet* is a complement to the E2 e.g.: *internet zaljubljenik* ('internet lover', i.e. one who loves the internet), *internet zavisnik*, *internet ovisnik* ('internet addict' i.e. one who is addicted to the internet), *internet dobavljač* ('internet provider' i.e. one who provides (somebody with) internet).

---

<sup>224</sup> A plausible explanation of the large percentage of domestic E2s (72 pct. in SrpKor and 67 in HNK) that combine with *auto* as E1 would be the fact that *auto* ends in -o, which is also considered a linking element in the traditional nominal compounds in Serbian and Croatian, e.g. *poljoprivreda* ('agriculture') and *pismonoša* ('letter bearer').

attributive endocentric

*internet sustav*

ATTR+HEAD

'internet system'

subordinate endocentric

*internet dobavljač*

COMPL+HEAD

'internet provider'

The other queried E1-nouns, listed in Table 4.4.6 are also found in both attributive and subordinate [N<sub>UNDECLN</sub>]<sub>N</sub>-compounds.

attributive endocentric

*auto-bomba* (SrpKor)

ATTR+HEAD

[N<sub>UNDECLN</sub>]<sub>N</sub>

'car bomb'

subordinate endocentric

*auto vožnja* (HNK)

COMPL+HEAD

[N<sub>UNDECLN</sub>]<sub>N</sub>

'driving of cars'

*parking usluga* (HNK)

ATTR+HEAD

'parking service'

*parking kontrolor* (SrpKor)

COMPL+HEAD

'parking attendant'

*kontakt tačka* (SrpKor)

ATTR+HEAD

'point of contact'

*kontakt improvizacija* (SrpKor)

COMPL+HEAD

'improvisation of contact'

*kamp mjesto* (HNK)

ATTR+HEAD

'camping spot'

*kamp lider* (SrpKor)

COMPL+HEAD

'camp site leader'

*business sektor* (HNK)

ATTR+HEAD

'business sector'

*biznis glavar* (SrpKor)

COMPL+HEAD

'chief of business'

The fact that both types of nominal compounds are formed shows us that this word-formational pattern competes not only with [A+N]<sub>NP</sub> as the recommended substitution for attributive compounds, e.g. *automobilska bomba*, *parkinška usluga* or *kontaktna tačka* but also with NPs with non-congruent complements [N+N<sub>obliq</sub>]<sub>NP</sub> as *vožnja autom* and *kontrolor parkinga*, glossed below.

[N+N<sub>obliq</sub>]<sub>NP</sub>

*vožnja autom*

HEAD+ COMPL

NOM + INSTR

'driving of cars'

[N+N<sub>obliq</sub>]<sub>NP</sub>

*kontrolor parkinga*

HEAD+ COMPL

NOM + GEN

'parking attendant'

So the replicated pattern which produces a nominal compound is productive, and it offers an alternative to adjectivisation or declension of the non-head in NPs.

*The nouns queried as element 2 in [N<sub>UNDECLN</sub>]<sub>N</sub>-compounds.*

Now, let us turn to the queried E2s and begin with an overview of how many types of undeclined E1-nouns the queried E2s combine with in each corpus.

Table 4.4.7 – Queried replicated E2-nouns in nominal compounds

E2-noun	no. of different E1-nouns		Translation
	SrpKor	HNK	
zona	22	5	zone
paket	17	6	package
šou/show	17	10	show
biznis/business	9	2	business
koktel/cocktail	1	0	cocktail (-party)
parket	0	2	parquet floor

Table 4.4.8 – Queried domestic E2-nouns in nominal compounds

E2-noun <sup>225</sup>	no. of different E1-nouns		Translation
	SrpKor	HNK	
muzika/glazba	32	27	music
čaj	3 (+6)	1	tea
sladoled	0	0	ice-cream

It is immediately evident from Tables 4.4.7 and 4.4.8 that the queried E2-nouns combine with a significantly lower number of E1-nouns than the queried E1-nouns above, in Table 4.4.6, combine with E2-nouns.

The domestic E2-nouns (Table 4.4.8) do not seem to be combined less or more with E1-nouns than the replicated E2-nouns (Table 4.4.7).

---

<sup>225</sup> Besides *sladoled* and *glazba* I have decided also to categorise *čaj* and *muzika* as domestic nouns because they are examples of the originally replicated nouns that can be considered to be *adoptees*, i.e. nouns which not even purist normativists think should be substituted with a domestic equivalent. The noun *muzika* is, however, considered foreign in Croatian, which is why I queried both *muzika* and *glazba* in both corpora.

There is, however, a rather large difference between Croatian (HNK) and Serbian (SrpKor) usage norms in terms of combinatorial possibilities, i.e. with how many different E1-nouns the queried E2-nouns are combined. In SrpKor *biznis/business* combines with 9 different E1-nouns, whereas only two different compounds are found in the HNK. The situation is similar regarding the E2s *paket* (SrpKor 17 : HNK 6) and *zona* (SrpKor 22 : HNK 5).

The relatively large number of compounds with either *muzika* or *glazba* are all compounds that denote a particular genre of music and all the E1-nouns, except one: *pleh* in *pleh-muzika* (brass wind music), are replicated from other languages, e.g. *hiphop* (hip hop), *džez* (jazz), *rok* (rock), *pop* (pop), *pačanga* (pachanga), *salsa* (salsa), *klecmer* (klezmer)

Compounds with *čaj*, *koktel/cocktail*, *sladoled* and *parket* as E2-nouns that meet the set criteria are so few, that we may conclude that compounds with these particular E2s have not entered the usage norm in the publicistic functional styles of Serbian and Croatian despite the fact that different scholars have pointed to them as examples of an innovative word-formational pattern in Serbian and Croatian, cf. above. The rare examples found are in HNK: *aroma čaj* (aromatic tea), *panel parket* ('click parquet flooring'), *ornament parket* ('ornamental parquet flooring'), and all three examples would have been excluded had they been product names, but they are included as they are types of a particular product and because the E1-nouns (*aroma*, *panel* and *ornament*) also exist in both languages as independent nouns. The examples found in SrpKor are: *rum koktel* ('cocktail with rum')<sup>226</sup> and *kleka čaj* ('juniper tea', i.e. tea from juniper berries) *filter čaj* ('bagged tea') *tonik čaj* ('tonic tea', i.e. a tea which is also a tonic). In Table 4.4.8, in a parenthesis, I have indicated that I found 6 more E1-nouns in compounds with *čaj* as E2-noun. They are all, however, only mentioned in a linguistic article within SrpKor as examples of incorrect naming of tea and can therefore not be considered representative of the usage norm of publicistic functional style of Serbian.<sup>227</sup>

The remaining four queried E2s: *zona*, *paket*, *šou/show* and *business/biznis* are all replicated E2s and I have only found two compounds where one of these E2s are combined with a Slavic domestic noun:

---

<sup>226</sup> Ćorić (2008) mentions *banana koktel*, and Starčević (2006) mentions *jagoda cocktail*.

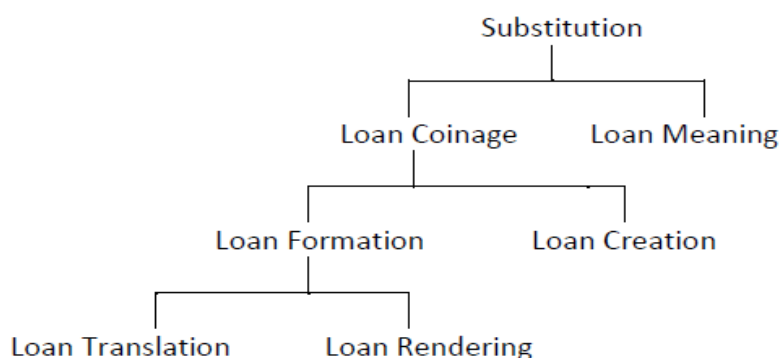
<sup>227</sup> These „incorrect“ compounds are: *hibiskus čaj* (hibiscus tea), *kamilica čaj* (camomile tea), *kopriva čaj* (nettle tea), *menta čaj* (mint tea), *šipak čaj* (dog-rose tea)

*poklon paket* ('gift package'), *Dunav zona* and two compounds where one of them is combined with an inherited noun *radio šou* ('radio show') and *vikend paket* ('weekend package deal' offered to guest at hotels, etc.)

The rest of the compounds with *zona*, *paket*, *šou/show* and *business/biznis* are all examples of compounds that comprise only replicated lexical matter. Some of them exist as compounds in the model language, e.g. *šou biznis* (show business), others as e.g. *internet paket* ('internet package') and *tampon zona* (< Fr. 'zone tampon' - 'buffer zone') are most likely the result of word-formation in the replica language according to the lexical pattern of English.

Applying Duckworth's categorisation of matter substitution (below and on p. 164) to the empirical data, we find examples of compounds, presumably replicated in their entirety (importation): *šou biznis* (< show business), and substitution in form of 'loan formation' through 'loan translation' of one of the elements: *internet dobavljač* (< internet provider), *kontakt tačka* (point of contact) and 'loan creation': *šoping špica* (< Eng. 'shopping' + Ger. 'Spitze')

Figure 4.1 - Duckworth's categorisation of matter substitution



Duckworth (1977: 40)

My analyses of the queried nouns, E1s and E2s, lead to the conclusion that only some of the innovations in word-formation of compounds in Serbian and Croatian proposed above have entered the usage norms of Croatian and Serbian, as documented in the two corpora. We primarily see compounds formed out of replicated lexical matter in this innovative way, i.e. Replicated matter in replicated lexical pattern, e.g. *šoping amater*, *šoping špica*, *šou biznis*, *tampon zona*, *business sektor* but also



many examples of Replicated and inherited matter in a replicated lexical pattern, e.g. *Internet paket, Dunav zona, kontakt točka, kamp mjesto, parking površina, salsa glazba, Internet dobavljač, šoping usluga, biznis glavar*. There are very few examples of Inherited matter in a replicated lexical pattern, and it is quite revealing, that some of these very few examples occur in a text where they are mentioned as examples of incorrect word-formation.

The fourth type of innovative compounds which I searched for were those formed out of entirely inherited matter but in a Replicated word-formational pattern, exemplified with *sladoled kocke* – ‘ice cream cubes’ in Table 4.4, p. 195. I found no examples of this type of compound in corpora.

My findings has led me to conclude that the innovative formation of nominal compounds in Serbian and Croatian explained and investigated in this study has indeed altered the structure of Serbian and Croatian publicistic usage norm but is generally reserved for nominal compounds in which at least one of the nouns is a piece of replicated matter.

In the *Cline e* below (also presented on p. 194) Serbian and Croatian usage norms are thus at “word-formational pattern” with the condition that one of the elements is a replicated noun.

*Cline e*

lexical >	>	>	>	>	> grammatical
content >	>	>	>	>	>function
matter >	>	matter + lexical pattern>	lexical pattern >	word-formational pattern	

The usage of [N<sub>UNDECL</sub>N]<sub>N</sub>-compounds in Serbian and Croatian is definitely present, despite the recommendations from the linguistic

The number of different E1- and E2-nouns which combine with the queried replicated nouns is larger in Serbian than in Croatian which I see as a reflection of the conclusions obtained in the analysis of replicated nouns and their recommended substitutions, which is that Croatian usage norm accepts less replicated lexical matter than Serbian usage norm. This in turn reflects the strong presence of purism in standard Croatian and its weaker presence in standard Serbian. Generally, puristic efforts target salient aspects of a language, i.e. matter rather than pattern. This being said, the users of the publicistic functional style of Croatian seem to be using the word-formational pattern [N<sub>UNDECL</sub>N]<sub>N</sub> almost as much as their Serbian counterparts. I will therefore speculate

that the reason behind the relatively lower number of these compounds in HNK is not due to reluctance in Croatian language users to use the replicated word-formational pattern but rather a reluctance to use the replicated matter involved.

In other words, Croatian and Serbian reactions to this replicated word-formational pattern are very similar which means that the strong normativist warnings against forming compounds in this manner are not followed by the users of Croatian and Serbian and thus a covert pattern change involving a rise from a minor to a major use pattern is underway in both Croatian and Serbian.

#### 4.5 Semiwords and affixes

In the previous section, focus was on the increased number of compounds consisting of two content items where both elements are free forms. We now turn our attention to content items which, like the nominal compounds, are the result of combining at least two elements but here one or more of these elements are a bound morpheme. I will subdivide these bound morphemes into the well-known category of affixes and the less known category of semiword. After a discussion of the formal and semantic features of affixes and semiwords I will propose a definition in form of a continuum which spans the dichotomies of concrete:abstract, specific:general and lexical: grammatical. Subsequently I will illustrate how some of the items have been lexicalised or grammaticalised in the replica languages. Finally, before the analysis of my empirical findings in HNK and SrpKor descriptions of semiwords and affixes in Serbian and Croatian will be presented.

The main works of reference I will be using on affixes and semiwords in Serbian and Croatian nouns are Klajn's *Tvorba reči u savremenom srpskom jeziku* ('Word-formation in contemporary Serbian') (2002; 2003), Babić's *Tvorba riječi u hrvatskom književnom jeziku* ('Word-formation in Croatian standard language') (1986), Ćorić's *Tvorba imenica u srpskom jeziku* ('Noun formation in Serbian') (2008) and Barić's *Imeničke složenice neprefiksane i nesufiksane tvorbe* ('Nominal compounds formed without suffixes and prefixes') (1980). I will also refer to the main grammars of Serbian (Piper and Klajn 2013) and Croatian (Barić et al. 1995) as well as various articles on individual phenomena in Croatian and Serbian. Finally, I will use works on word-formation in general: Scalise and Bisetto's "The Classification of compounds" in *The Oxford Handbook of Compounding* (2011) and Bauer's *Introduction to Morphology* as well as "Is there a class of neoclassical compounds, and if so is it productive?" (Bauer 1988; 1998).

In the investigation of nouns focus was on frequency and proportionality between the use of replicated nouns and their proposed substitutes. In the investigation of nominal compounds lexical units were still involved, but focus was rather on their combinatorial possibilities which, in turn, served to show how entrenched the possibility to form [N<sub>UNDECL</sub>N]<sub>N</sub>-compounds had become. In the investigation of semiwords and affixes focus will stay on combinatorial possibilities in order to show the productivity

of replicated lexical and grammatical formants in the two languages. And, because the investigated linguistic units represent both lexical and grammatical matter focus will also be on this difference and whether their level of entrenchment in Serbian and Croatian usage norms is dependent on this difference. The focus has thus, motivated by the objects of investigation, moved from content and lexicality towards function and grammaticality.

*Examples of semiwords and affixes*

<b>papir-o-logija</b>	'excessive amounts of paper, red tape'
<b>raket-o-drom</b>	'spaceport'
<b>ćuto-log</b>	'silent, introvert person'.
<b>pop-art</b>	'pop art'
<b>art-prevara</b>	'art scam'
<b>akva-park</b>	'water park'
<b>auto-portret</b>	'self portrait'
<b>avio-služba</b>	'flight service'
<b>pseudo-nauka</b>	'pseudoscience'
<b>vele-sila</b>	'super power'
<b>jugo-nostalgija</b>	'Yugo-nostalgia' <sup>228</sup>
<b>etno-kuća</b>	'ethnic house, ethnic home'
<b>eko-škola</b>	'eco-school'
<b>elektro-napon</b>	'electrical current'
<b>skribo-man</b>	'manic writer'
<b>aero-drom</b>	'airport'

---

<b>ko-voditelj</b>	'co-anchor', 'co-host'
<b>re-izbor</b>	'reappointment'
<b>bicikl-ijada</b>	'cycling contest'
<b>pit-erija</b>	'pita restaurant'
<b>tenis-er</b>	'tennis player'
<b>nato-izacija</b>	'complying with NATO-standards'
<b>kugl-ana</b>	'bowling alley'

Below the dotted line all the emphasised elements are considered to be affixes; above the line we find affix-like elements which in Serbian and

---

<sup>228</sup> „Yugonostalgia can be broadly defined as nostalgia for the fantasies associated with a country, the SFRY, which existed from 1945 to 1991” (Lindstrom 2003: 233).

Croatian tradition are called *afiksoidi* subsuming *prefiksoidi* and *su-fiksoidi* ('affixoids', 'prefixoids' and 'suffixoids') (Klajn 2002; Košutar and Tafrá 2009; Horvat and Štebih-Golub 2010) or *vezani leksički morfem* ('bound lexical morpheme') (Barić 1980), and which have recently also been described as *nesamostalni formant* ('unautonomous formant') (Slijepčević 2013) and finally, inspired by Russian linguistics *vezana osnova i okrnjena osnova* ('bound stem' and 'clipped stem') (Jovanović 2013). Among experts in other languages they have been termed (Russian) 'radixoid' and 'terminoelement' (Dragičević 2009), 'unifix', 'uniradixoid', 'lexicomorpheme', 'adjectival morpheme', 'semiprefix', 'quasiprefix', 'preposed blocks', 'agglutinative element', 'analytical adjective', 'premorpheme' (Ćorić 2008) and (English) 'semiwords' (Scalise and Bisetto 2011) or simply 'stem' (Bauer 1998).<sup>229</sup> I use the term *semiword*, because a) it is an English term, b) it is a term that covers bound lexical morphemes regardless of their position<sup>230</sup> and c) the term indicates that it is in many ways similar to a word (a free form) without being a full word.

In English, *semiword* or *stem* is used in the description of neoclassical compounds, i.e. compounds formed in a modern language using Classical Greek and Latin nominal stems as for instance:

<i>geo-logy</i>	- study of earth
<i>photo-graph</i>	- drawing made by light (Bauer 1998: 405)

Bauer alerts us to the fact that lexemes, which are formed out of a neoclassical stem and a full lexeme, carry the same semantic relationship between its elements as a neoclassical compound does, exemplifying it with nouns as:

<i>eco-doomster</i>	- person who foretells doom in ecological matters
<i>hamburger-ology</i>	- study of hamburgers

Bauer subsequently calls the result of these combinations a „compromise between the neoclassical nature of neoclassical compounds and the native nature of native [English] compounds“ (Ibid: 407). He then concludes, on the basis of further analysis „...that neoclassical compounding

---

<sup>229</sup> Some of the terms listed here also cover the undeclined preposed adjective or noun which are otherwise free morphemes, i.e. the E1 in the nominal compounds, described. Other terms cover either only preposed or postposed semiwords.

<sup>230</sup> in the same way as the term 'affix' covers a number of different grammatical morphemes including prefixes and suffixes

acts as some kind of prototype, from which actual formants may diverge in unpredictable ways.“ (Ibid: 409)

So, the realisation that semiwords belong to a specific category, neoclassical or not, which combine with full lexemes or other semiwords in compound-like structures is not new, nor does it pertain only to Slavic.

That semiwords are primarily international and have entered the European languages from the classical languages Greek and Latin is also emphasised by Serbian and Croatian researchers. However, they also underline the fact that semiwords are not necessarily of classical origin; they may also have been replicated from other languages and even semiwords of domestic origin exist (Barić 1980: 77), (Klajn 2002: 157).

Bound morphemes such as *-ijada*, *-er*, *-erija*, *-izam*, *-logija*, *-man*, *-manija*, *-fil*, *-ijada*, *-drom*, *-holičar*, *hidro-*, *akva-*, *ko-* *re-* have all been replicated as part of foreign lexemes, and need not even have been perceived as morphemes at the time of replication. But, as the number of replicated lexemes ending in, e.g. *-logija* increases, the ending is perceived as having its own meaning, its own content. This is in itself not so intriguing, as most of these morphemes have the same meaning in the model language, for instance: *-logija* – KNOWLEDGE OF/SCIENCE OF. From our perspective, however, it becomes increasingly interesting when these morphemes combine with new lexical stems, i.e. when a speaker of the replica language creates innovations by forming new words with such morphemes. In these instances we may talk about replication of grammatical matter or rather; the reanalysis of a string of sounds from being a bound lexical morpheme in specific replicated lexical matter (words) to being piece of grammatical matter, but still a bound morpheme, with which new words may be formed in the replica language.

In the literature on this topic, it is recognised that this is the general way in which replicated bound morphemes are integrated into the replica language. In the *Handbook of Language Contact* (Hickey 2010) we learn that a prerequisite for the word-formational extension of a specific bound morpheme seems to be a large amount of replicated lexical matter, i.e. words which include the same bound morpheme (Winford 2010: 176). In this way the morpheme, which may or may not be a derivational affix or a semiword in the model language, becomes just that in the replica language.

As mentioned in section 4.2, replicated matter may extend its semantic scope in the replica language, independently of its semantic scope in the model language. In Serbian and Croatian, the semantic scope of the

bound morpheme *-logija* which at some point was first replicated, from Latin and/or Classical Greek in words such as e.g. *teologija* 'theology' has, for instance, been extended to also mean 'excessiveness' in *papirologija* ('excessive amounts of paper, red tape')(Jovanović 2013: 302).

#### **4.5.1 Formal and semantic features of semiwords and affixes**

Because I consider the distinction between grammatical and lexical matter and pattern important to the overall argument of what is more or less replicable as well as what has a more or less likelihood of being accepted by the linguistic authorities, it is important to distinguish between semiwords and derivational affixes in terms of grammaticality and lexuality. The defining features of semiwords are described formally (morphologically and word-formationally) and semantically.

##### ***Formal features***

- Semiwords share with affixes the ability to combine with a series of lexemes (Klajn 2002: 144).
- Semiwords are bound lexical stems whereas lexemes are free lexical stems (Scalise 1986: 75)
- If a preposed element exists as the stem of other, by suffixation derived, wordforms, the preposed element is undoubtedly a semiword and not a prefix, because it is impossible to derive a word by applying a suffix directly to a prefix (Klajn 2002: 145).
- Semiwords are, as are affixes, formally (morphologically) unautonomous: 'the main difference between lexemes and semiwords lies in the fact that a lexeme exists autonomously whereas a semiword does not exist autonomously' (Barić 1980: 17)<sup>231</sup>
- Semiwords may be clipped lexical stems or full, but bound, lexical stems.

##### ***Semantic features***

- Semiwords share with lexical stems (mostly nominal stems) a unique lexical meaning. (Klajn 2002: 144) (Mihaljević 2012: 72) (Barić 1980: 16)
- Semiwords have "...a semantic value or density more similar to that of lexemes than to that of many prefixes." (Bauer 1998: 407)

---

<sup>231</sup> Original: „Osnovna je razlika između pune riječi (leksema) i vezanog leksičkog morfema, dakle, u tome što puna riječ dolazi samostalno, a vezani leksički morfem ne dolazi samostalno...”

- Semiwords do not, as affixes do, modify the lexeme they are combined with, but concretize it (Barić 1980: 17)<sup>232</sup>.
- Prefixes are equivalent to function words such as prepositions, whereas preposed semiwords are equivalent to content words such as adjectives and nouns, which means that *polu-* and *vele-* are semiwords because *polu-* is an allomorph of the adverb *pola* – 'half' and *vele-* an allomorph of the adjective *veliki* – 'big', whereas *protu-* and *anti-* are prefixes because they are the semantic equivalent of the preposition *protiv* – 'against', and finally *pro-*, and *super-* are prefixes because they are the semantic equivalents of the prepositions *kroz* – 'through' and *iznad* – 'above, over'. (Babić 1986: 23).

Despite the different naming and exemplification of these elements, the definition is relatively clear and places the semiwords formally as one would place *affixes* (which are *function items*) but semantically as *content items* (equivalent to lexemes). In other words, semiwords share morphological and word-formational features with affixes, but semantic features with nouns, adjectives and adverbs.

Semiwords are, as we have now seen, not the same as derivational affixes, but they are not free forms (lexemes) either. So let us briefly look at some definitions of affixes and lexemes.

- Affixes are obligatorily bound morphs. Derivational affixes produce a new lexeme from a stem. Through adding a derivational affix, the lexical category of the stem may be changed, which is often the case with derivational suffixation. A prototypical derivational affix does not have a fixed meaning, that is, it does not change the meaning of a stem in the same way every time it is applied to a stem. (Bauer 1988: 12–13; 86)
- Suffixes may never be stems, they are never used independently, they only have meaning when they are attached to a stem and they cannot be the semantic core in a lexeme. Suffixes merely modify the meaning of the lexical stem (Barić 1980: 21)
- Lexeme is traditionally a term used to cover all possible grammatical forms of one word. All lexemes belong to a word category.

---

<sup>232</sup> Original: „Međutim, ti glasovni nizovi ne modificiraju značenje riječi nego ga konkretiziraju.“. I have not attempted to make an idiomatic translation of “konkretiziraju”, as I must surmise that Barić means that the preposed semiword makes the meaning of the resulting noun less abstract, narrowing the scope of possible referents.



In other words, semiwords are *content items* and as such *lexical matter* with the morphological restrictions of an affix, which is a *function item* and *grammatical matter*. It is therefore, according to the hierarchies introduced in section 3.3.2, to be expected that semiwords are replicated more easily than grammatical matter, such as true affixes but still not as easily as free content items, such as nouns, adjectives and adverbs.

#### 4.5.2 Semiwords

Compounds with semiwords (sw) have existed in Serbian and Croatian at least since the mid-19<sup>th</sup> century and were, as the [N<sub>UNDECL</sub>N]<sub>N</sub>-compounds ignored and demoted by linguists in Serbia and Croatia for the better part of the 20<sup>th</sup> century (Klajn 2002: 141). Preposed semiwords are, however, today the most efficient individual means of creating new terms and the number of compounds with preposed semiwords is constantly rising (Ibid: 148). Many of the semiword-compounds in Serbian and Croatian are neoclassical compounds, that is, they consist of two classical content items, of which at least one is a semiword as for instance: *eko-logija* ('ecology') or *aero-drom* ('aeropot'), but it has been observed that the number of semiword-compounds with inherited elements (underlined) is increasing as in: *avio-sluz̑ba* ('flight service'), *jugo-nostalgija* ('Yugo nostalgia'), *akva-park* ('water land') *etno-kuća* ('ethnic house') and *ćuto-log* ('silent, introvert person').

The number of nouns consisting of one or more semiwords is increasing in Croatian and Serbian and this increase is seen as an effect of globalisation by both Croatian (Horvat and Štebih-Golub 2010: 2) and Serbian scholars (Ćorić 2008: 117). Ćorić also informs us, referring to Gutschmidt (2001)<sup>233</sup>, that this tendency towards internationalisation has been particularly strong in the former socialist, currently transitional countries. In recently published literature which treats this phenomenon, nouns with preposed semiwords are often commented and analysed alongside [N<sub>UNDECL</sub>N]<sub>N</sub>-compounds<sup>234</sup> as both types of nouns predominantly consist of a preposed undeclined replicated element (E1) and a noun (replicated or not) as the second element (E2), and the E1 typically modifies, determines or, in Barić's words, concretizes the E2 which is also the head of the [sw N]<sub>N</sub>-compound, e.g. *akva-park* ('water land').

---

<sup>233</sup> Гутшмидт, Карл (2001): Тенденция интернационализации в современных славянских литературных языках – отражения поверхностные и глубинные, in: *Internacionalizácia v súčasných slovanských jazykoch: za i proti*, Bratislava, 7-19

<sup>234</sup> by e.g. Slijepčević (2013), Ćorić (2008) and Horvat and Štebih-Golub (2010)

Nouns consisting of two semiwords [SW SW]<sub>N</sub>, as e.g. *aero-drom* ('airport') also exist. In these instances the semanto-syntactic structure is the same as in [SW N]<sub>N</sub>-compounds, that is, the E1 modifies or determines the E2. We may therefore label the semantic structure of these compounds as 'attributive' and 'subordinate' but not 'coordinate' as there are no semiword compounds in which both elements are, in a semantic sense, heads as is characteristic of coordinate compounds (cf. p. 187).<sup>235</sup>

In Serbian and Croatian, there are much fewer postposed semiwords than preposed ones. Barić (1980: 69) mentions 12-14 and Klajn (2002: 160-63) mentions 20 as the most frequent ones. Postposed semiwords are always as far as I can surmise head of the compound. The E1 in such compounds is a lexical stem (verbal, nominal or adjectival), and it only exists as a modifying attribute so postposed semiwords may only be heads in compounds and these compounds will always be 'attributive' compounds as defined in *Attributive compounds*, p. 188ff. Thus we have the following three possible models of nominal semiword compounds:

<i>akva-park</i>	<i>aero-drom</i>	<i>bicikl-ijada</i>
[SW <sub>mod/det/conc</sub> N] <sub>N</sub>	[SW <sub>mod</sub> SW] <sub>N</sub>	[STEM <sub>mod</sub> SW] <sub>N</sub>
'waterpark'	'airport'	'cycling contest'

There is, however, a problem with the definition that delimits postposed semiwords from postposed derivational affixes, i.e. suffixes. It has, as far as I know, never been claimed that suffixes such as *-ište/-lište* or *-ona/-onica* are postposed semiwords. They do, however, seem to meet the criteria for being semiwords, that is, they do not have, as the prototypical affix, an inconsistent, abstract, meaning, or as Bauer would put it: *-ište/-lište* or *-ona/-onica* do change the meaning of a stem in the same way every time they are applied to a stem (Bauer 1988: 12-13; 86) because 'Derivatives ending in the suffix *-ište* derived from nouns most often denote a place, and rarely anything else' (Babić 1986: 123)<sup>236</sup> According to Babić this is also true for a number of suffixes ending in *-ište*, e.g. *-lište*, *-ilište*, *-elište* (Ibid: 125-126) 'Derivatives ending in the suffix *-onica* denote a room, in which an action is done, rarely anything else' (Ibid: 161)<sup>237</sup> 'Through the suffix *-ona*, nouns are formed out of verbal stems

<sup>235</sup> Exemplified by *boy-actor* and *pjesnik-pjevač* ('singer-songwriter')

<sup>236</sup> Original: „Izvedenice sa sufiksom *-ište* izvedene od imenica najčešće označuju mjesto, a rijetko što drugo (stvar)“

<sup>237</sup> Original: „Izvedenice sa sufiksom *-onica* znače prostoriju u kojoj se obavlja glagolska radnja, rjeđe što drugo“

and they denote a room or a building.' (Ibid: 318)<sup>238</sup>. There are several other suffixes in Serbian and Croatian, which have more or less unchanged meanings. If we were to apply Babić's criteria for distinguishing between prefixes and preposed semiwords along with Bauer's definition of a prototypical affix (cf. p. 225 and 220) suffixes with a constant meaning such as *-ište* - PLACE would qualify as a semiword because of its semantic equivalence to the noun *mjesto* - 'place'.

Another problem in the formal and semantic descriptions of semiwords and affixes is that one could argue that the domestic lexical stems (underlined) in nominal compounds such as *rukopis* - 'manuscript, handwriting' and *nogomet* - 'football' should in fact be categorised as semiwords because they do not exist autonomously. However, it is not my aim to redefine word-formation in Croatian and Serbian, so I will not pursue this notion further, but let it stand as an illustration of the weaknesses of the existing definitions of what a semiword is; a weakness that stems from the fact that semiwords are primarily associated with replicated matter and only secondarily used to label domestic matter. Elements such as *-pis* and *-met* are, when mentioned in the literature, referred to as lexical bound stems with a Ø-suffix, by, for instance, Klajn (2002: 53).<sup>239</sup>

In my opinion, the deciding factor when determining whether a bound morpheme is a derivational affix or a semiword should be the degree of specificity of the denotatum of the element. The more specific the denotatum of the bound morpheme is, the closer it is semantically to being a postposed semiword and not a suffix. Barić (1980: 22) makes the same distinction when she explains the difference between the suffix *-ac* in *starac* ('an old man') and the semiwords *-fil* and *-fob* in *rusofil* ('Russophile') and *rusofob* ('Russophobe') as being one of formational [function] vs. lexical meaning [content] and says that *-ac* means simply 'man' whereas *-fil* and *-fob*, means 'man who loves' and 'man who hates'<sup>240</sup>. Klajn (2002: 144–51) presents various attempts at delimiting affixes from semiwords but comes to the conclusion that it is not possible to

---

<sup>238</sup> Original: „Sufiksom *-ona* tvore se imenice od glagolskih osnova i znače prostoriju ili zgradu.“

<sup>239</sup> Stjepan Babić does not mention them at all in his monograph on word-formation (Babić 1986), cited in this section. Bauer (1998) on and Scalise (1986), referred to in subsection 4.5.1 above also refer to semiwords as *stems*.

<sup>240</sup> Original: „Element *-fil* u *rusofil* ne znači samo 'čovjeka' kao što *-ac* u *stàrac* nego 'čovjeka koji voli', kao što i *-fob* znači 'čovjeka koji mrzi'.“

identify any one feature which could serve as a criterion for distinguishing between all semiwords and affixes.

Because a strict delimitation of affixes and semiwords is very difficult to find, I suggest defining them as points along a continuum which spans the dichotomies of *concrete:abstract* and *specific:general* which correspond to the continuum used in grammaticalisation theories with the opposites: *lexical (content)* vs. *grammatical (function)*. Below I have placed 8 examples of Serbian and Croatian nouns formed by combining two elements, of which either E1 or E2 is a noun, in order to describe the gradual difference in terms of *content* and *function*.

Figure 4.3 - A continuum from lexeme to affix<sup>241</sup>

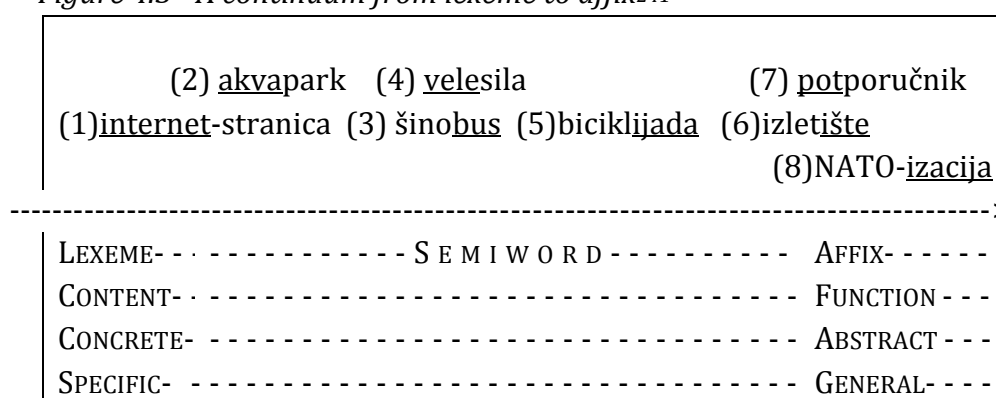


Figure 4.3 shows a continuum which at its leftmost end has an example of composition (1), where a *content* item, which refers to a *concrete* entity, represented by a *lexeme* (the noun *internet*) which carries a *specific meaning* is combined with another noun (*stranica*). Further on along the continuum I have placed examples (2-4) of what in the literature is called *semiwords*<sup>242</sup>, beginning with the most *concrete, content*-ful preposed lexical item with *specific meaning* (2) (*akva*-WATER) followed by (3) the clipped nominal stem *-bus* – VEHICLE 'whose purpose or type is determined by the preposed element' (Barić 1980: 70), but carries *lexical* and *content*-ful but *less specific* and *less concrete* meaning. In (4) the preposed adjectival element *vele*- BIG is placed after *-bus* because it carries a *more abstract, more functional* and *more general* meaning than *-bus*. Using Barić's terms *vele*- modifies more than it concretizes the postposed noun (*сила*). Example (5) *-ijada* has been identified as a semiword by

<sup>241</sup> 1) Internet page, 2) water park, 3) street car, 4) super power, 5) cycling contest, 6) excursion site, 7) 2nd lieutenant, 8) complying with NATO-standards

<sup>242</sup> Or the equivalent of semiwords

Ćupić (2002: 2: 79)<sup>243</sup> but as a suffix by Ćorić (2008: 53–62). The fact that *-ijada* has only one rather specific meaning, e.g. a competitory event as in *somijada* – 'catfish-catching contest' or *univerzijada* – 'an international sporting competition among students from different places of higher learning' or *biciklijada* – 'bicycling contest' makes it possible to categorise it as a *semiword* with the *specific, lexical content*: COMPETITIVE EVENT. Should *-ijada* take on a more abstract meaning such as EVENT of any kind and thus combine more freely with free lexical stems it should be recategorised as a *suffix* as *-ište* in (6) *izletište*, which has the rather *general, abstract* meaning PLACE.

Were this abstraction of the meaning of *-ijada* to take place, it would, indeed be an instance of *grammaticalisation* as it would be a change from being less grammatical to more grammatical or, in lexicalisation terms: a *delexicalisation* of *-ijada*.

*pot-* in (7) *potporučnik* – '2nd lieutenant' (*pot-* - 'sub'/'under') is a classic example of a prefix, which *functions* as modifier of its stem *poručnik* – 'lieutenant'. It does not denote any *specific* entity but the *general, abstracted* spatial relation 'under'. Finally, we have *-izacija* in (8) *Natoizacija* which here exemplifies how *suffixes function* as a derivational means to derive a new lexeme with a meaning and/or word category distinct from that of its lexical stem, and the *semantic content* of the suffix is *abstract* and *general*.

### ***Semiwords and affixes in the lexical and grammatical clines***

The uncertain categorisation of items as either semiwords or affixes is, as I have shown, inextricably connected to the dichotomy of lexicality and grammaticality.

Semiwords and affixes, which have entered Serbian and/or Croatian through replication, are replicated as part of one or more lexemes. They may or may not be free morphemes in the model language. They may or may not be considered morphemes in the model language. Over time, the replicated matter may, become a more free formant in the replica language, i.e. combine with several lexical stems as *art-* in *art-kafić* ('art café'), *art-glazba*, ('art(istic) music') *art-mješavina* ('art mixture'), *art-turizam* ('art tourism'), *akva* in *akvapark* ('water land'), *akvaterapija* ('water therapy/aqua therapy') and *-ijada* in *kobasicijada* ('sausage con-

---

<sup>243</sup> Ćupić uses the terms suffixoid and quazisuffix with an autonomous meaning



mantic bleaching and has been extended to combine with domestic nominal stems as *mačkolog*<sup>246</sup> and verbal stems as in *ćutolog*<sup>247</sup> and reanalysed to mean PERSON who is in some emphatic relation to the denotatum of E1. The relation between the denotatum of E1 and the person (E2) is not as specific as in the internationalisms and this is the result of the semiword *-log* taking on more abstract and less fixed meaning, which is, as mentioned above, a defining feature of affixes.

#### *Preposed semiwords in Serbian and Croatian*

In Serbian and Croatian the number of existing preposed semiwords is much larger than that of the postposed ones. Nominal compounds consisting of a preposed semiword and a noun [SW N]<sub>N</sub> are very similar to [N<sub>UNDECL</sub>N]<sub>N</sub>-compounds, the main difference being that the E1<sub>sw</sub> does not exist as an independent (free) morpheme.

Whereas the existence of domestic postposed semiwords is rejected in the literature, some domestic preposed semiwords are recognised. Among them we find *vele-* GREAT an allomorph of the adjective *veliki* – 'big' and *jugo-* a clipped form of the adjectival stem *jugoslavensk-/jugoslovensk-* ('Yugoslav') or the nominal stem *jugoslavij-* ('Yugoslavia').

As this investigation of semiwords and affixes in Croatian and Serbian is aimed towards contact-induced innovations where the innovation lies in the word-formation in Croatian and Serbian and where replicated semiwords and affixes are combined with lexical stems in the replica language, many of the examples of semiword-compounds in Barić (1980) and Klajn (2002) are of less interest because they are examples of nouns replicated in their entirety, nouns which, unsurprisingly, consist of replicated semiwords and affixes. Concretely, replicated nouns such as *termostat*, *fotografija*, *fonologija*, *rehabilitacija*, *milimetar*, *kilogram*, *uniseks*, *hiperinflacija*, *etnolog*, *mikroorganizam*, *antifašizam*, *koegzistencija*, *homofobija*<sup>248</sup> all consisting of semiwords and or affixes of foreign origin are of no interest in this part of the study. Some of the bound morphemes, however, may be interesting, because they are also present in words formed in Serbian and Croatian as e.g. *kovoditelj*,

---

<sup>246</sup> *mačka* - cat

<sup>247</sup> *ćut(j)eti* – to stay silent

<sup>248</sup> In English: thermostat, photography, phonology, rehabilitation, millimetre, kilogram, unisex, hyperinflation, ethnologist, microorganism, antifascism, inn, coexistence, homophobia

*etnokuća, papirologija, ćutolog, kuglana, reizbor, NATO-izacija* (see p. 221 for translations).

Ćorić (2008: 120) describes the model of these compound formations in Serbian, including nominal compounds as: 'a component of foreign origin + an independent lexeme as the base'<sup>249</sup> and adds that the number of lexemes formed in this manner has grown rapidly and involves an almost unlimited number of not only common nouns but also proper nouns. (Ibid:). Ćorić provides an extensive list of preposed elements, alphabetically ordered, which at the time of writing (2008), he believed would give a 'pretty realistic picture of this expansive type of noun formation'. Ćorić, however, does not distinguish between preposed nouns, semiwords and prefixes in this list of 293 preposed replicated elements (Ćorić 2008: 126–49). Ćorić also provides a selected list of more than 700 compound-like lexemes in Serbian in which the E1 (prefix, semiword or lexeme) is replicated and the E2 is a domestic or inherited (previously replicated and adopted) lexeme.

In 1980, Barić (1980: 78–82) provided a list of 96 preposed semiwords of foreign origin and 9 of Croatian origin. She also gives a complete list of nominal compounds in Croatian which do not contain any affixes, i.e. a list of all nominal compounds, which consist of two independent lexemes (as the nominal compound: *duhankesa* cf. p. 201) or where one or two of the elements are semiwords (as in *avioslužba* and *papirologija*). In 2002, Klajn provides a list of 93 of the most frequent preposed semiwords of foreign origin in Serbian of which all but one are replicated form either Greek or Latin. (Klajn 2002: 152–56)

Even though the lists provided by Klajn and Barić are focused on preposed semiwords alone, they do not, as does Ćorić, emphasise and exemplify the ones that are seen in compounds with domestic lexemes.

In order to determine which preposed semiwords I should include in my searches within the electronic corpora I have compared the lists provided by Ćorić (2008) with the lists provided by Barić (1980) and Klajn (2002). Furthermore, I have searched the normative works (language manuals, the published decisions from the Croatian *Council* and Serbian *Board* on standardisation issues) to ascertain whether the linguistic authorities of Croatian and Serbian have made any normative

---

<sup>249</sup> Original: „komponenta stranog porekla + samostalna leksema kao osnovinska reč“



recommendations concerning the use of specific semiwords. On this basis I have chosen to search the corpora for the following preposed semiwords:

*aero-, akva-, art-, auto-, avio-, bruto-, eko-, etno-, jugo-*

#### *Postposed semiwords in Serbian and Croatian*

As mentioned above, in Serbian and Croatian there are much fewer postposed semiwords than preposed ones. Both Barić (1980: 69) and Klajn (2002: 160–63) list only classical (Greek and Latin) semiwords<sup>250</sup> and reject the existence of domestic postposed semiwords. Barić (1980: 68; 78–82) and Barić et al. (1999: 245–46) list the semiwords in Croatian and Ćorić (2008: 154–57) and Klajn (2002: 151–56) do the same for Serbian. Horvat and Štebih-Golub(2010) list what they consider the most recent and most frequently used semiwords in Croatian and Jovanović (2013), Slijepčević (2013) and Piper and Klajn (2013: 230) do the same for Serbian. On the basis of their and own observations, I have chosen to search the corpora for the following postposed semiwords:

*-art, -bus, -holičar, -holik, -log, -logija, -man, -manija, -teka*

#### **4.5.3 Derivational affixes**

In Serbian and Croatian, affixes are used widely in word-formation and are thus very useful in the ever needed replenishment of lexis, through derivation. Foreign derivational affixes enter, in the same way as foreign semiwords, a language first as part of replicated lexemes and only then might they become independent, but bound, morphemes, i.e. semiwords or affixes with which new lexemes are composed or derived in the replica language. Derivational affixes are however, as opposed to semiwords, *function* items which serve to derive new lexemes out of existing ones, e.g. *reizbor* ('reappointment'), *ultrazvuk* ('ultrasound') *piterija* ('pita restaurant'), *teniser* ('tennis player'). That derivational affixes are viewed as function items does not mean that they are without content. The content of derivational affixes is, as illustrated in Figure 4.3, above, only more abstract and less specific than that of lexical items such as semiwords. The

---

<sup>250</sup> Klajn also mentions the existence of the English postposed semiword *-gejt* ('-gate'), but discards examples of its use in Serbian, in e.g. *Dafinagejt* ('Dafinagate', a embezzlement scandal related to bank-owner Dafina Milanović,) as 'typical journalistic invention which in addition never lose the traits of proper nouns, so therefore we cannot consider them as genuine novelties in the Serbian vocabulary' (Klajn 2002: 164)

delimitation of affixes and semiwords is not always clear and sometimes an item can be categorised as both being a semiword due to its concrete and specific meaning, e.g.: *ijada* – COMPETITIVE EVENT or as an affix due to its productivity as a derivational tool.<sup>251</sup>

The number of replicated semiwords, especially preposed ones is, as mentioned above, large and rising in both Serbian and Croatian. However, quite a large number of replicated affixes are also part of Croatian and Serbian. I have again chosen to focus primarily upon nouns which have been derived with a replicated affix, so as to make the results comparable with the results from investigating innovations in nouns, nominal compounds and nominal semiword compounds.

Many of these affixes only combine with replicated lexical stems and exist in the same combination in many other languages so it may be difficult to determine whether the affixation happened in the model language, in which case it is the whole lexeme that has been replicated (cf. p. 167), or the affixation happened in the replica language, in which case the affix has been recognised by the speakers as an affix and used in derivation of new lexemes. In other words, the speakers of the replica language reanalyse a string of phonemes which occur in a row of replicated lexemes (content items) as an affix (function item) and begin deriving new lexemes using this affix. This is, for instance, the case with the suffix *-ist(a)* which recurs in a multitude of replicated lexemes, e.g. *komunist(a)*, *socijalist(a)*, *optimist(a)* but also exist a number of lexemes derived in the replica language through affixation with *-ist(a)*; e.g.: *šahist(a)* ('chess player'), *taksist(a)* ('taxi-cab driver') and *vezist(a)* ('signal man').

Klajn (2002: 196) informs us that the vast majority of the replicated affixes are of Greek and Latin origin (as was the case with the semiwords). Furthermore, he informs us that the international prefixes are mostly seen on nouns, less on adjectives and even less so on verbs, whereas domestic prefixes combine most easily with verbs, less with adjectives and even less so with nouns (Ibid: 194). Replicated matter is in other words, more present in nouns than in other word categories, not only when it comes to lexical matter in simplex nouns as *marketing*, nominal compounds as *basket parket* or semiword compounds as *akvaterapija* but also when it comes to the grammatical matter of affixes as in *vezist(a)* ('signal man') or *reizbor* ('reappointment').

---

<sup>251</sup> More than 50 different nouns derived with *ijada* in Serbian and in Croatian were found in HNK and SrpKor.

A large number of the replicated affixes in Serbian and Croatian are not new and have existed in the languages for a long time. As the focus here is on contact-induced innovations in the last 20-30 years, I will concentrate on a) the few affixes that in the literature have been recognised as recent innovations and/or b) especially productive grammatical formatives in recent years and/or c) on those replicated affixes that have been singled out as markers of either Croatian or Serbian.

The affixes under investigation here are either posed before or after the lexical stem, i.e. only prefixes and suffixes are investigated. Bearing in mind the lack of a clear distinction between semiwords and affixes in the literature, we will now look upon what Serb and Croat linguists consider to be replicated affixes.

#### *Replicated suffixes in Serbian and Croatian*

Ćorić (2008) concludes that there are two genetically defined groups of replicated suffixes, international and Turkish. International: **-ija**, *-acija*, *-encija*, *-icija*, **-ist(a)**, *-alist(a)*, *-onist(a)*, *-enist(a)*, **-at**, *-ijat*, *-orat*, *-onat*, **-or**, *-tor*, *-ator*, *-itor*, **-ant**, **-ent**, **-(it)et**, **-ik**, **-ura**, **-esa**, **-iz(a)m**, **-ada**, **-al**. Turkish: **-džija**, *-adžija*, *-edžija*, **-lija**, *-alija*, **-luk**, *-aluk*, **-ana**. (Ćorić 2008: 51). The suffixes in bold are primary suffixes, whereas the immediately subsequent ones are variations of them. Ćorić also singles out *-acija*, *-džija*, *-luk*, *-ist(a)*, *-at*, *-ator* and *-ant* as the most productive ones (Ibid:), and later also points to the Turkish suffixes *-lija* and *-ana* as being productive in Serbian (Ibid: 195). He does not, however, comment on whether they are frequently combined with domestic or replicated stems. In his description and exemplification of the individual suffixes, there are no examples of domestic stems with the most productive international suffixes (*-acija*, *-ist(a)*, *-at*, *-ator* and *-ant*). With the exception of *-at* (which is also combined with Turkish stems) Ćorić's examples suggest that these suffixes only combine with international stems (Ibid: 44-49). However, in the description of one of the two most productive suffixes replicated from Turkish (*-luk*) it is highlighted that it has entered the Serbian word-formational system on a grand scale (Ibid: 47), because it is, as are *-lija* and *-ana*, used to derive lexemes with a stem which is either non-Turkish or even domestic.

Ćorić singles out *-ist(a)*, *-izam*, *-ijada* and *-izacija* as very productive suffixes not only as suffixes on replicated stems but also as grammatical formatives of domestic stems (Ibid: 157). Other scholars have pointed to the fact that the replicated suffixes *-er* and *-erija* have become more productive in recent years. (Muhvić-Dimanovski 2005) and (Kekez 2012).

In recent grammars of Croatian (Barić et al. 1995) and Serbian (Piper and Klajn 2013) the origin of the suffixes is either not mentioned at all (Piper and Klajn 2013) or only mentioned in the description of some Turkish suffixes (-*džija* and -*lija*) but not of the international (e.g. -*ist*) nor of the Turkish suffix -*ana* (Barić et al. 1995). In both grammars, the section on nominal suffixation, i.e. nouns derived by suffixation, emphasises that not all suffixes will be mentioned. Piper and Klajn (Piper and Klajn 2013: 222) claim to list only the most frequent ones<sup>252</sup>, Barić et al. (1995: 305) claim to list the strongly productive, the weakly productive and one or two unproductive suffixes<sup>253</sup>. Neither of the grammars mentions the suffixes -*acija*, -*luk*, -*at*, -*ator* and -*ant*, which Ćorić (2008) singled out as the most productive replicated suffixes alongside -*ist(a)* and -*džija*. The two normative grammars also leave out the international suffixes -*izam*, -*izacija*, -*erija*, -*er* and -(*ij*)*ada*. Whether it is truly the case that these suffixes are unproductive or less productive or it is an example of (covert) purism is difficult to say. In the recommendations by the Croatian council and the Serbian board on standardisation, suffixation is hardly mentioned at all. When contact-induced linguistic innovations are concerned, these institutions have focused more on lexical matter (replicated nouns, adjectives and compounds) and matters of spelling (of replicated nouns and replicated compounds). cf. sections: 2.4.3, p.103ff, 4.2.1, p.159ff and 4.4.2 p.203ff.

Babić (1986) provides a list of no less than 506 suffixes in noun formation (Babić 1986: 50–51). Babić does not categorise replicated suffixes separately but does, occasionally, inform his readers that a particular suffix combines with replicated stems. But Babić's examples represent, in fact, the result of replication of whole nouns and not the replication of the individual suffixes. For instance, he exemplifies the suffix -*ada*, with nouns where -*ada* is not a suffix, but part of the lexical stem, such as *marinada* – 'marinade', *blokada* – 'blockade', etc. (ibid. 79). Klajn (2003), on the other hand, devotes a separate 35-page long section to replicated (foreign) nominal suffixes<sup>254</sup>, excluding the Turkish -*luk*, -*lija*, -*džija* and -*ana*, as he views them as so adopted that they should be treated as domestic or inherited suffixes, rather than foreign ones. (Klajn 2003: 218)

---

<sup>252</sup> Original: „Od nekoliko stotina imeničkih sufiksa, ovde se može ukazati samo na nekoliko desetina najčešćih.“

<sup>253</sup> Original: „Navode se samo vrlo plodni i slabo plodni, a od neplodnih tek pokoji.“

<sup>254</sup> “Именички суфикси - страни” (Klajn 2003: 218–53)

On the basis of the claims in the literature, which I have referred to above, on nouns derived in Croatian and Serbian of lexical stems with replicated suffixes, I have chosen to search SrpKor and HNK for the following suffixes:

- *izam* as in *životinjizam* ('animal-like behaviour') derived from the domestic lexical stem *životinj-* of the noun *životinja* - 'animal'.
- *ijada* as in *somijada* ('catfish-catching competition'), derived from the domestic lexical stem *som-* of the noun *soma* - 'catfish'.
- *erija* as in *piterija* ('pie place'), derived from the inherited<sup>255</sup> lexical stem *pit-* of the noun *pita* - 'pie'.
- *er* as in *kombajner* ('combine driver'), derived from the replicated lexical stem *kombajn-* of the noun *kombajn* - 'a combine'.
- *ant* as in *prevarant* ('con-man') derived from the domestic lexical stem *prevar-* of the noun *prevara* - 'fraud'.
- *ator* as in *gnjavator* ('pestilent person') derived from the domestic lexical stem *gnjav-* of the verb *gnjaviti* - 'to pester'.

#### *Replicated prefixes in Serbian and Croatian*

The replicated (or foreign) prefixes are singled out as a specific category in the normative works on Serbian and Croatian. Babić (1986) mentions 19, Klajn (2002) 31. They are all replicated from the classical languages, except for *des/dez-* which is from French 'dés-')

The following 14 prefixes are according to Klajn (2002) and sometimes Babić (1986) productive formants in combination with domestic nouns:

- anti-* as in *antirodoljub* ('anti-patriot')
- eks-* as in *ekskralj* ('ex-king', i.e. former king)
- ekstra-* as in *ekstrazarada* ('extra earnings')
- ko-* as in *kopredsednik* ('co-president')
- kontra-* as in *kontranapad* ('counterattack')
- kvazi-* as in *kvazi-umetnik* ('quasi artist')
- maksi-* as in *maksisuknja* ('maximum size skirt')
- mini-* as in *minikuhinja* ('mini kitchen', 'kitchenette')
- para-* as in *paraknjiževnost* ('paraliterature')
- pseudo-* as in *pseudorešenje* ('pseudo-solution')
- re-* as in *reizbor* ('reappointment')

---

<sup>255</sup> *pita* is an earlier replicated noun from Turkish, to which it was, it is believed, replicated from modern Greek (Kekez 2012: 54)

*super-* as in *superradnik* ('super-worker')  
*ultra-* as in *ultradesničar* ('ultra rightwing politician')  
*vice-* as in *vicekralj* (viceroys)  
*hiper-* as in *hiperveza* ('hyper-connection')

The fifteenth entry on the list, *hiper-* is not labelled a prefix that combines with domestic nouns by neither Klajn nor Babić but both Slijepčević (2013: 325) and Otašević (1998: 101) mention *hiper-* as a preposed element which may be combined with domestic nouns. Out of these 15 prefixes I have chosen to search the SrpKor and HNK for:

*hiper-, eks-, kontra-, ko-, kvazi-, para-, re-*

### ***Normative attitudes towards replicated semiwords and affixes***

Whereas there was a clear difference between the prescribed norms of standard Serbian and Croatian as regards the use of replicated lexical items (nouns and compounds), the differences in the prescribed norms concerning semiwords and suffixes are smaller. Even though the Croatian linguistic authorities recommend 'substitution of all replicated prefixes', even when combined with a replicated lexeme, they discourage substitution of suffixes and semiwords of Latin or Greek origin as 'they have no practical substitutes' (Barić et al. 1999: 204, 246; IHJJ 2015b; 2015d)<sup>256</sup>.

The Serbian language guide „Srpski jezički priručnik“ (P. Ivić et al. 2011) holds no specific recommendations regarding the use of foreign prefixes but encourages the use of replicated matter, especially in professional terminology, and tells its readers that 'it is good that the majority of civilized peoples have the same words for the same phenomena', that it is more natural to use replicated matter in the exact sciences and that it would lead to chaos if, for instance, all chemical terms were to be substituted with domestic creations (P. Ivić et al. 2011: 61)<sup>257</sup>.

---

<sup>256</sup> Original (Barić et al. 1999: 204): „Tuđe prefikse valja, gdje god je to moguće zamijeniti domaćima: *anti-* s *protu-*, *a-* s *ne-*, *trans-* s *preko-*, *pseudo-* i *kvazi-* s *nadri-* ili *nazovi-*.“ Original (Ibid: 246): „Nemaju praktičnu zamjenu.“

<sup>257</sup> Original: „...да је добро што за поједне појмове већина језика цивилизованих народа има исту реч. [...]народчито у егзактним наукама, у техници и медицини,[...] природније (је) употребљавати стране речи. Није тешко замислити хаос који би настао ако бисмо покушали да све хемијске термине заменимо домаћим кованицама.“

So, as regards semiwords and suffixes of Greek and Latin origin the recommendations of Croatian and Serbian normativist are quite similar (although for different reasons).

#### 4.5.4 Empirical data

Inspired by the lists of semiwords and affixes provided by scholars and referred to above, I have chosen to query HNK and SrpKor for the following semiwords and affixes:

<u>Preposed semiwords/Prefixes</u>	<u>Postposed semiwords/Suffixes</u>
art-	-art
akva-	-bus
jugo-	-teka
aero-	-holičar
avio-	-holik
eko-	-manija
etno-	-man
bruto-	-logija
hiper-	-log
eks-/ex-	-ijada
kontra-	-ator/-itor
kvazi-	-er
para-	-erija
re-	-ant
ko-/co-	-izam

I have placed the items in these lists corresponding to their respective position in the continuum presented as “ Figure 4.3 - A continuum from lexeme to affix”, p. 229. At the top of the lists we find the items with the most concrete and specific meaning and at the bottom of the lists we find the items with the least specific and most abstract meanings. Furthermore, the items at the top of the list have more lexical content than those at the bottom.

#### *Search criteria*

All the queries in both corpora were compiled in concordances which I inspected in order to find the nouns representing the following innovative word-formations:

a) [**domestic stem + semiword**]<sub>N</sub>, e.g. *igroteka* in:

*Prva zagrebačka i hrvatska **igroteka** u kojoj će djeca s posebnim potrebama...* (HNK)

(‘The first *collection of children’s toys* where children with special needs will...’)

Domestic stem: *igr-* from *igra* (‘game’) or *igrati* (‘to play’)

b) [**domestic stem + replicated suffix**]<sub>N</sub>, e.g. *zajebant* in:

*..samo da bi provocirao i izigravao kontraša – **zajebanta*** (SrpKor)

(‘..only so you could pretend to be a oppositionist – a kidder’)

Domestic stem: *zajeb-* from *zajebati* (‘to screw over’)

c) [**replicated prefix + domestic noun**]<sub>N</sub>, e.g. *reizgradnja* in:

*Suština se, pak, vidi u **reizgradnji** međusobnog poverenja* (SrpKor)

(‘The essence, however, is visible in the *rebuilding* of mutual trust’)

Domestic noun: *izgradnja* (‘building up, construction’)

d) [**semiword + domestic noun**]<sub>N</sub>, e.g. *aerozagađenje* in:

*Ne jenjava **aerozagađenje** u Pančevu* (SrpKor)

(‘The *air pollution* in Pančevo is not subsiding’)

Domestic noun: *zagađenje* (‘pollution, act of polluting’)

In other words, I wished to ascertain the level of productivity in word-formation with replicated bound grammatical matter (affixes) and with replicated bound lexical matter (semiwords) in combination with domestic lexical matter.

However, I also wished to ascertain the productivity of word-formations, (with the queried semiwords and affixes) which do not hold any domestic matter but may still be considered formed in the replica language, i.e.

e) [**replicated prefix/semiword + replicated noun/semiword**]<sub>N</sub>, e.g.:  
*avio-bomba* (‘aerial bomb’), *eks-premijer* (‘former prime minister’),

f) [**replicated stem/semiword + replicated suffix/semiword**]<sub>N</sub>, e.g.:  
*snobizam* (‘snobbishness’, ‘snobbery’), *papirologija*, (‘excessive amounts of paper’, ‘red tape’) *bibliobus* (‘library bus’), *programoteka* (‘collection of computer programmes’, ‘programme file’)



Concretely, I singled out the nouns in which the word-formation was likely to have occurred in the replica language but as I do not know the word-formational history of each noun, the relevance of the selected results is based on my estimation. These criteria have led me to exclude nouns such as *snouborder*, *roker*, *nokauter* (snow-boarder, rocker, knock-outer), *šinobus*, *pop-art*, *skriboman*, *kleptoman*, *kinoteka* ('bus on rails, pop art, manic writer, films archive') which have certainly or very probably been replicated in their entirety, i.e. they are examples of replicated nouns which happen to contain the queried post-posed elements. The same criteria led me to include *čokoladoman*, *asfalter*, *autostoper*, *bibliobus* and *filmoteka* ('chocolate craze, asphalt paver, hitch-hiker, library bus, movie collection') as they are examples of nouns which have a high probability of having been formed in the replica languages, despite the fact that they consist entirely of replicated matter. Finally, I also conducted searches of one domestic semiword: *jugo* to determine the productivity of this relatively new semiword, formed through clipping of the adjectival stem *jugoslavensk-/jugoslovensk-* ('Yugoslav') or the nominal stem *jugoslavij-* ('Yugoslavia').

#### *Preposed elements*

In the ensuing Table 4.5 of preposed elements (semiwords and prefixes), I have for both corpora indicated how many different elements each of the queried semiwords and affixes combines with. I did, however, not take into account the wordforms which are names of products, businesses, companies, festivals, events, TV-programmes etc. on the same grounds as I omitted these findings in the samples of nominal compounds.

Table 4.5 – Queried preposed semiwords and prefixes<sup>258</sup>

Preposed semiwords	no. of different E2s			
	SrpKor	HNK		
art	66	57		
akva/aqua	7	5		
jugo	68	141		
aero	26	18		
avio	114	53		
eko	121	204		
etno	133	89		
bruto	62	68		
Prefixes	SrpKor	HNK	Recommended substi- tutes in HJS <sub>259</sub>	
hiper	57	31		-
eks/ex	30	53		<i>bivši</i> <sub>260</sub> , <i>raz</i>
kontra	142	87		<i>protu</i>
kvazi	156	129		<i>laži, nadri, nazovi</i>
para	36	34		<i>laži, nadri, nazovi, ne, pa, polu</i>
re	134	141		<i>ponovni</i>
ko/(co) <sub>261</sub>	53	25		<i>su-</i>

With the exception of *akva/aqua* we see that word-formation with the selected preposed semiwords and prefixes are, in quite widespread use in the two corpora. I view the productivity and quantity of nouns formed with the selected preposed elements as indicative of the general usage of these elements in Croatian and Serbian. A large quantitative difference in the number of E2s between the usage norms of Serbian and Croatian as documented in the corpora is presumably dependent on the individual preposed element and not on the willingness of the

<sup>258</sup> For a complete list of the findings behind the figures in Table 4.5, cf. Appendix B.

<sup>259</sup> The Croatian Language Manual (*Hrvatski jezični savjetnik*) Barić et al (1999) recommends substitutes for the replicated prefixes, as listed here. *hiper-* is not included as a prefix in the HJS.

<sup>260</sup> Except the adjectives *bivši* ('former') and *ponovni* ('repeated'), all the recommended substitutions are prefixes

<sup>261</sup> There were no nouns, meeting the criteria, in either corpus, with the prefix *ko-* spelled *co-*.

language users to derive or accept and implement nouns with these preposed elements.

On the one hand we have *avio*, *etno*, *hiper* and *kontra* which are around 80-100 pct. more productive as formant E1s in SrpKor. On the other hand we have *jugo*, *eko* and *eks/ex* which are 80-100 pct. more productive as formant E1s in the HNK. The rest of the queried preposed elements are used in approximately the same amount of types of nouns in both corpora. I interpret this as indicative of the equal acceptance and implementation in both languages despite the different recommendations concerning replicated matter in general and specifically the Croatian recommendations to avoid the use of replicated prefixes.

The following lists of findings serve to illustrate the diversity of the nouns with the listed preposed semiwords or prefixes. The queried semiwords and prefixes are found in the corpora combining freely with domestic nouns ( $N_{\text{dom}}$ ) and replicated nouns ( $N_{\text{rep}}$ ) forming nouns such as the following:

<b>[PFX+N<sub>dom</sub>]<sub>N</sub></b>	<b>Translation</b>	<b>Source</b>
<i>eksvladika</i>	‘ex-bishop’	(SrpKor)
<i>exuposlenik</i>	‘ex-employee’	(HNK)
<i>kontratužba</i>	‘counteraccusation’	(HNK)
<i>kontrazahtev</i>	‘counter-demand’	(SrpKor)
<i>kopredlagač</i>	‘co-proposer’	(SrpKor)
<i>kopredsjedatelj</i>	‘co-chairman’	(HNK)
<i>kvazioporbenjak</i>	‘quasi-opposition member’	(HNK)
<i>kvazivoda</i>	‘quasi-leader’	(SrpKor)
<i>paraliječnik</i>	‘paramedic’	(HNK)
<i>paravlada</i>	‘para-government’	(SrpKor)
<i>reimenovanje</i>	‘renaming’	(HNK)
<i>reobrazovanje</i>	‘re-education’	(SrpKor)

<b>[PFX+N<sub>rep</sub>]N</b>	Translation	Source
<i>koministar</i>	'co-minister'	(SrpKor)
<i>koselektor</i>	'co-manager of national team'	(HNK)
<i>kontraimidž</i>	'counter image'	(SrpKor)
<i>kontravolej</i>	'counter hit (in tennis)'	(HNK)
<i>kvaziombudsman</i>	'quasi-ombudsman'	(SrpKor)
<i>paražurnalizam</i>	'quasi-journalism'	(HNK)
<i>eks-teritorija</i>	'former territory'	(SrpKor)

<b>[SW+N<sub>dom</sub>]N</b>	Translation	Source
<i>aeroprostor</i>	'air space'	(SrpKor)
<i>aerotvrtka</i>	'airline company'	(HNK)
<i>akva promet</i>	'water traffic'	(HNK)
<i>art delo</i>	'work of art'	(SrpKor)
<i>art-vodič</i>	'art guide'	(HNK)
<i>avio stajanka</i>	'holding bay'	(HNK)
<i>avio-gorivo</i>	'airplane fuel'	(SrpKor)
<i>bruto cena</i>	'gross price'	(SrpKor)
<i>bruto satnica</i>	'gross payment per hour'	(HNK)
<i>eko dažbina</i>	'eco-tax'	(SrpKor)
<i>ekonovinar</i>	'organic reporter'	(HNK)
<i>etno-nasleđe</i>	'ethno heritage'	(SrpKor)
<i>etno pokret</i>	'ethno movement'	(HNK)
<i>hiperpijaca</i>	'hyper-market'	(SrpKor)
<i>hiper-uspješnost</i>	'hyper success'	(HNK)
<i>jugodopisnik</i>	'Yugoslav correspondent'	(HNK)
<i>jugo zvezda</i>	'Yugoslav star'	(SrpKor)

<b>[SW+N<sub>rep</sub>]N</b>	Translation	Source
<i>aerobaza</i>	'air base'	(SrpKor)
<i>art scena</i>	'(the) art scene'	(HNK)
<i>aviolinija</i>	'airline route'	(SrpKor)
<i>eko-zona</i>	'organic zone'	(HNK)
<i>hipernormiranost</i>	'excessive normation'	(HNK)

As was the case with the nominal compounds I found examples of semiword-compounds where the  $E1_{sw}$  (the preposed semiword) was not an attribute to the head of the compound, but a complement, e.g.:

<i>aerozagađenje</i>	<i>art vodič</i>
COMPL+HEAD	COMPL+HEAD
'air pollution'	'art guide'

In both semiword compounds above the non-head is not a descriptive attribute of the head, but rather a complement to the head, i.e. the air is being polluted (by someone) and the guide is guiding (someone) through the art.

### *Postposed elements*

Because of the vast number of wordforms of nouns ending in some of the selected postposed elements, the searches of the corpora were not absolute in the same manner as the searches of preposed elements were.<sup>262</sup> On the other hand, I found that some of the searched postposed elements are scarcely productive or not productive at all in the usage norm of Serbian and Croatian as documented in the two electronic text corpora. Among these scarcely productive elements we find *-art*, *-bus*, *-holik*, *-holičar* and *-itor*.

*-art*, *-bus*, *-holik*, *-holičar* and *-itor*

Neither of these postposed elements seems to be productive in word-formations in the replica languages. I did not find any nouns with the postposed semiword *-itor*, which might have been formed in Serbian and Croatian. In other words, *-itor* is not according to my findings a productive suffix in Serbian and Croatian usage.

The lexical item *art* may be a linguistic item that many speakers of Serbian and Croatian understand and use as a preposed semiword to form new nouns, but as a postposed semiword it is mostly if not only used in combination with other replicated items in nouns which were most likely formed in the model language, as e.g. *bodi art* ('body art'), *sajber-art* ('cyber art'), *popart* ('pop art') (SrpKor) *folk art* ('folk art'), *mail-art* ('mail art'), *videoart* ('video art') (HNK) In other words, *art* may act as a postposed semiword, but there is very little evidence in the

---

<sup>262</sup> See appendix C for tables of the word types which were selected from the concordances and included in the analysis.

two corpora to support the idea that nouns are formed by compounding a noun with *art* as E2 in Croatian and Serbian<sup>263</sup>.

Concretely, I found only one example of *art* as a postposed semiword in a noun formed in the replica language: *rat-art* ('war art') (HNK) and only two examples of *-bus*: *bibliobus* ('library bus') (both corpora) and *konjobus* ('horse drawn bus') (SrpKor). *rat-art* and *konjobus* consist of a domestic nominal stem (*konj* ('horse') and *rat* ('war')) followed by the semiword (*-art* and *-bus*), viz. [STEM<sub>dom</sub> + SW]<sub>N</sub> whereas *bibliobus* consists of two replicated semiwords *biblio-* (< *biblioteka* - 'library') and *-bus* (< *autobus* - 'bus'), viz. [SW + SW]<sub>N</sub>.

*-holik* and *-holičar* ('-holic', i.e. person addicted to the denotatum of E1) are in use in both corpora, *-holičar* being the more adapted version (i.e. a clipping of *alkoholičar* 'person addicted to alcohol') whereas *-holik* is the clipped stem of the English nominalised adjective *alcoholic*, i.e. *-holic*. The semiwords *-holičar* and *-holik* are used synonymously, are *nomina agentis*, and bear the meaning: PERSON ADDICTED TO OR PERSON BEING EXCESSIVE regarding the E1 of the noun.

The postposed semiwords *-holik* and *-holičar* are, as *-art*, *-bus* and *-itor*, in very scarce use in word-formation in Croatian and Serbian as documented in HNK (5 types of E1) and SrpKor (6 types of E1) but *-holik* and *-holičar* do form nouns in the replica languages with both replicated stems (STEM<sub>rep</sub>) and domestic stems (STEM<sub>dom</sub>):

---

<sup>263</sup> Not including company names and the like.

**[STEM<sub>dom</sub> + SW]<sub>N</sub>**

*kupoholičar* (SrpKor) (*kup-* < *kupiti* – ‘to buy’, *kupovina* – ‘shopping’)

*dućanoholičar* (HNK) (*dućan* – ‘shop’)

*radoholik, radoholičar* (HNK, SrpKor) (*rad-* < *raditi* – ‘to work’)

**[STEM<sub>rep</sub> + SW]<sub>N</sub>**

*čokoladoholik* (HNK),

*čokoholičar* (SrpKor) (*čokolad-* and *čok-* < *čokolada* – ‘chocolate’)

*blogoholičar* (SrpKor) (*blog* – ‘(we)blog’)

*netoholičar* (SrpKor) (*net* < *internet* – ‘The Internet’)

*šopingoholičar* (HNK) (*šoping* – ‘shopping’)

*seksoholičar* (HNK) (*seks* – ‘sex’)

Some of these findings are easily recognised as calques of English nouns such as shopaholic, workaholic, chocoholic, blogoholic, etc. Nevertheless, they are formed in the replicated language, sometimes with domestic stems, so the English semiword ‘-holic’ has been replicated into Serbian and Croatian and may be considered a formant of new lexemes in the usage norm of both languages.

-teka, -manija, -man, -logija, -log, -ijada, -ator, -erija, -er, -ant and -izam

The rest of the queried postposed semiwords and suffixes are productive as formants of nouns in both replica languages as represented in the two corpora. Except in the case of *-erija* and *-er* I have found nouns derived from a domestic stem as E1, viz. [STEM<sub>dom</sub> + SW]<sub>N</sub> and [STEM<sub>dom</sub> + SFX]<sub>N</sub>.

**[STEM<sub>dom</sub> + SW/SFX]<sub>N</sub>**

-teka:

*željoteka* (*želja* – ‘wish’) – place where your wishes come true (SrpKor)

*igroteka* (*igra* – ‘game’, *igračka* – ‘toy’) – place to play, a collection of toys (HNK)

*biljoteka* (*bilje* – ‘flora’ and *apoteka* – ‘pharmacy’) – pharmacy with herbal medicine (SrpKor)

-manija:

*baštomanija* (*bašta* – ‘garden’) – ‘garden-mania’ (SrpKor)

*srbomanija* (Srbin – ‘(a) Serb’) – ‘focus on all things Serb’ (HNK)

*sajmomanija* (*sajam* – (a) fair) – ‘fair mania’ (SrpKor)

*zv(j)ezdomanija* (*zv(ij)ezda* – ‘star’) – ‘celebrity mania’ (SrpKor, HNK)

-man:

*glumoman* (*gluma* – '(the art of) acting') – 'person crazy about acting' (SrpKor)

*kavoman* (*kava* – 'coffee') – 'coffee-maniac',  
ie. 'person obsessed with coffee' (HNK)

-logija:

*ćuftologija* (*ćufta* – 'meatball') – 'expertise in meatball eating' (SrpKor)

*čvorologija* (*čvor* – 'knot') – 'expertise in tying knots' (SrpKor)

*lupetologija* (*lupetati* – 'to talk nonsense deliberately') –  
'nonsensical talk' (HNK)

*licologija* (*lice* – '(a) face') – 'science of facial expressions' (HNK)

-log:

*ćutolog* (*ćuteti* – 'to stay silent') – 'quiet, introvert person' (SrpKor)

*nogometolog* (*nogomet* – 'football') – 'football expert' (HNK)

-ijada:

*kobasičijada* (*kobasica* – 'sausage') – 'sausage competition'  
(SrpKor, HNK)

*roštiljijada* (*roštilj* – 'barbecue') – 'barbecuing event' (SrpKor)

-ator:

*gnjavator* (*gnjaviti* – 'to pester') – 'pestilent person' (SrpKor, HNK)

*snagator* (*snaga* – 'force') – 'a very strong person' (SrpKor, HNK)

-ant:

*jurišant* (*juriš* – '(mil.) attack') – 'person with a nervous affliction, due to his/her war time experiences' (SrpKor)

*zabušant* (*zabušati* – 'to avoid doing one's job, to skulk') – 'person who is avoiding to do his/her job, a skulk' (SrpKor, HNK)

-izam:

*pučizam* (*puč* – 'coup') – 'coup-making' (HNK)

*uništizam* (*uništiti* – 'to destroy') – 'inclination to destroy' (HNK)



[STEM<sub>rep</sub> + SW/SFX]<sub>N</sub>

-er:

*tračer* (trač (< Ger. Tratsch) – ‘gossip’) –  
‘gossiper, i.e. person inclined to gossip’ (SrpKor, HNK)

-erija:

*koketerija* (*koketirati* (< Ger. kokettieren) – ‘to flirt’) – ‘flirtatious activity, flirting’ (SrpKor, HNK)

For further examples, cf. appendix C.

The fact that most of these postposed elements are combined with not only replicated but also domestic stems makes it evident that these replicas are entrenched not only in the language systems of Croatian and Serbian, but also in the usage norm of their speakers. In other words, it is not only possible to use these replicated postposed elements when forming words in Croatian and Serbian, their use is also accepted by the language users regardless of normative recommendations.

### ***Summing up the semiwords and affixes***

In this section we have seen how semiwords and affixes may not always be exactly delineated, but rather should be described along a continuum from lexeme to affix. Whereas derivation of nouns from lexical stems with affixes is recognised as a common and largely productive word-formational strategy in Serbian and Croatian, the combination of a semiword with a lexical stem is regarded as a replicated word-formational strategy in the same way as compounding of nouns in the [N<sub>UNDECL</sub>N]<sub>N</sub>-pattern is.

In spite of the Croatian normativists’ negative attitude towards semiwords replicated from languages other than Classical Greek and Latin, and towards affixes replicated from any language, the results of the queries in the corpora suggest that the usage norm of Croatian is in no way in agreement with the prescribed norm. Furthermore, the results indicate that Serbian and Croatian usage norms are very similar when it comes to replication and implementation of semiwords and derivational affixes alike.

As the reader will recall, I consider Serbian and Croatian structurally identical. So, the fact that the publicistic writers of Croatian and of Serbian use derivational affixes as well as semiwords in a similar manner tells us a number of things.

First, it shows us that the usage norms of the two languages are more aligned concerning the use of replicated semiwords and the grammatical matter of affixes than they are when the lexical matter of replicated nouns and replicated nominal compounds is concerned. This may be seen as an indication that the structural sameness of the two languages increases the likelihood that they will implement the same contact-induced innovations as both speech communities are under the same global influences.

Secondly, the similar Croatian and Serbian usages of replicated grammatical matter and dissimilar usages of lexical matter make it evident that it is more difficult to convince language users to substitute the bound morphemes and grammatical matter of affixes and semiwords than it is to convince them to substitute lexical matter (be it an independent noun or an undeclined E1 in a nominal compound)

Thirdly, the same grammaticality or degree of entrenchment in the linguistic system and usage norm of the replica language is most likely the reason for the resilience of the replicated affixes against the substitutions recommended by the Croatian linguistic authorities. In other words: The more grammatical a contact-induced linguistic innovation is, the lower the probability of its 'successful' substitution with a domestic equivalent will be. Its grammaticality makes it a more entrenched part of the replica language and it is therefore harder to diminish its use, i.e. make it unacceptable and thus not part of the norm.

#### 4.6 Indefinite articles

After having searched and researched *replicated lexical matter* (nouns and semiwords) *replicated grammatical matter* (affixes) and after having searched and researched a *replicated lexical pattern* and the possibility of a *replicated word-formational pattern* (both in nominal compounding) we will now move one step further along the path from the concrete, specific and lexical towards the abstract, general and functional by looking into the possible emergence of indefinite articles in the otherwise article-less languages: Serbian and Croatian.

As opposed to the phenomena under investigation in the previous sections which all include replicated matter either directly as in *kvaliteta*, *džez muzika*, *šoping špica*, *internet veza*, *kupoholičar*, *reobrazovanje* or indirectly, i.e. relayed pattern replication, as in as in *sladoled kocke* and *nana čaj*, the phenomenon under investigation in this section involves no replicated matter at all. We have, in other words, now proceeded to *replication of a grammatical pattern*. By replicating a grammatical pattern the structure of the replica language is altered. Where replication of matter leads to innovation of the building blocks, and replication of word-formational patterns leads to innovation in how to arrange these building blocks within the architectural design, the replication of a grammatical pattern changes the architectural design itself. When the architectural design is changed without replication of any matter, it follows that existent matter is used in new functions. When a piece of existent lexical or grammatical matter takes on a new *function* it is termed *grammaticalisation* (cf. also p. 119ff)

Linguistic innovations induced by grammatical pattern replication, resulting in grammaticalisation in the replica language, have much more profound implications for the linguistic system in a language than replication of matter. Such innovations are less likely to receive the attention of purist linguists. The reason for this lack of attention is at least three-fold. First, the innovations are not noticed because the matter is not foreign but domestic. Secondly, such innovations are much fewer in number, and thirdly, it takes much longer for a pattern replication to be accepted, actualised and implemented by ‘cohorts of speakers’ so the process of change is much more gradual which makes it less detectable. It is, using the term proposed by Gottlieb (2006), *covert*.

The need to name a new concrete phenomenon as e.g. *aerodrom* (‘airport’) is sudden and will be met quickly (through matter replication or lexical pattern replication [calquing]). In contrast, the tendency to

mark something as indefinite by existent linguistic means (in a way similar to that of other languages) is neither sudden nor urgently necessary. The motivation behind replication of a grammatical pattern is, in other words, more elusive.<sup>264</sup> Not only is it more difficult to ascertain the speakers' general motivation for innovating their grammar, it is also difficult to make evident that it is indeed a contact-induced innovation and if so to identify the Model-language.

The need among the Serbian and Croatian speakers to mark the referent of an NP as either 'definite' or 'indefinite' by way of an article might stem from language contact, and the replication of such a grammatical pattern, may therefore have occurred because of the intense contact with a language or languages that contain this pattern. Up until the beginning of the 20<sup>th</sup> century, some of the article-languages of contact which were used as languages of administration or religion in the Serbian and Croatian language area were German and Italian in the north and west, Greek and Turkish in the east and south part. It is however, as mentioned above, difficult to present evidence to prove that a specific grammatical pattern is indeed replicated. The probability that a grammatical innovation in a language is contact-induced is, however, higher when matter has been replicated from a language or languages that have exactly this pattern. Sarah Thomason, whose *Borrowing scale* was introduced in section 3.3, describes a situation where the contact is realised through a small group of reasonably fluent bilingual speakers of the replica language who are nevertheless only a minority among its speakers. In this situation it is likely that content items and even function items are replicated, but there will only be a few instances of replication of grammatical patterns which could involve new functions and new functional restrictions for existing syntactic structures, or increased usage of previously rare word orders. (Thomason 2001: 70). The usage of *jedan* and *neki* as indefinite articles could easily be interpreted as an instance of a "new function for an existing syntactic structure" in Serbian and Croatian.

---

<sup>264</sup> The motivation for replication of word-formational patterns is also less straightforward, but, as argued earlier, the presence of replicated matter with an inherent replicated pattern is likely to act as a conductor through which the pattern is relayed from replicated to domestic matter.

Heine and Kuteva (2010) support this assumption even more by arguing that the grammaticalisation of articles in Slavic languages is contact-induced<sup>265</sup> because the Slavic languages „that exhibit at least minor use patterns of articles [are] exactly those languages that are known to have a long history of contact with article languages such as German, Italian, or Greek“ (Heine and Kuteva 2010: 91).

Even if historical linguists were to make evident that article-like use of *jedan* and *neki* is not contact-induced but rather a language-internally motivated grammaticalisation, their increased use as indefinite articles may still be ascribed to contact. Heine and Kuteva (2005) call such an occurrence “a rise from a minor to a major [grammatical] use pattern”. As mentioned in section 3.3.1, Heine and Kuteva characterise such a rise as follows:

*The rise of a major use pattern in contact-induced replication*

- a. An existing use pattern is used more frequently
  - b. It is used in new contexts
  - c. It may become associated with a new grammatical function
- (Heine and Kuteva 2005: 45)

So, in a study of contact-induced innovations in Serbian and Croatian the use of *jedan* and *neki* are interesting because:

- a) the use of *jedan* and *neki* has been described as “article-like” and signalling indefiniteness while being part of languages which are generally seen as article-less
- b) *jedan* and *neki* have been identified as being contact-induced by scholars over a 100 years ago (Maretić 1899)<sup>266</sup> and recently (Radovanović 2009; Heine and Kuteva 2006).
- c) the use *jedan* and *neki* is, according to some scholars, increasing. (Radovanović 2009; 2003; 2000; M. Ivić 1971; Katunar et al. 2013).

In the following description of indefinite article-like usage of *jedan* and *neki* in Serbian and Croatian I will among other sources refer to Lyons, C. (1999), who examines definiteness in general, Heine and Kuteva (2006 and 2010) who examine the emergence of definite and indefinite articles, Breu (2012) who examines the possible existence of an

---

<sup>265</sup> Heine and Kuteva term it “replica grammaticalisation” (Heine and Kuteva 2010: 90)

<sup>266</sup> See p. 138 for quote.

indefinite article in two Slavic languages in intense contact with the article-languages German and Italian, Hauge (1977), who examines the use of the Bulgarian equivalent to *jedan*: ‘един’, Katunar et al. (2013) who examine specificity markers in Croatian, and finally Milka Ivić (1971) who examines the usage of *jedan* and *neki* in Serbo-Croatian. But, first let us see how *jedan* and *neki* are described formally and functionally in normative and educational works as Stevanović (1964), Barić et al (1995), Piper and Klajn (2013), Mrazović and Vukadinović (1990), Alexander (2006), Mønnesland (2002) and Barjamović-Orahovski (2006).

### ***Normative descriptions of jedan and neki***

The word class of the two items in focus are numeral (a content item) and pronoun (a function item), respectively. They may be translatable as ‘one’, and ‘some’. Morphologically they are declined in gender, number and case, here exemplified by *jedan*<sup>267</sup>.

	<u>Singular</u>		
	<u>masculine</u>	<u>neuter</u>	<u>feminine</u>
nom	jedan-Ø	jedn-o	jedn-a
acc	jedan-Ø/jedn-og <sup>268</sup>	jedn-o	jedn-u
gen	jedn-og	jedn-og	jedn-e
dat	jedn-om	jedn-om	jedn-oj
instr	jedn-im	jedn-im	jedn-om
loc	jedn-om	jedn-om	jedn-oj
	<u>Plural</u>		
	<u>masculine</u>	<u>neuter</u>	<u>feminine</u>
nom	jedn-i	jedn-a	jedn-e
acc	jedn-e	jedn-a	jedn-e
gen	jedn-ih	jedn-ih	jedn-ih
dat	jedn-im	jedn-im	jedn-im
instr	jedn-im	jedn-im	jedn-im
loc	jedn-im	jedn-im	jedn-im

Syntactically, *jedan* and *neki* are determiners in NPs.:

*jedno staro selo*

*neka lijepa djevojka*

<sup>267</sup> The desinences in the declension of *neki* differs only in the nom.sing.masc.: *nek-i*

<sup>268</sup> If the referent is inanimate, *jedan* is used, if animate *jednog*.

‘a/one old village’

‘some/a pretty girl’

In the normative descriptions of the non-numerical usage of *jedan*, it is often stated that it signals indefiniteness and that it is used in a manner similar or identical to the usage of the indefinite pronoun *neki*:

‘The numeral *jedan* is in our language very often used – not to mark a number, but rather as a kind of indefinite article. [...] By using the word *jedan*, in reality one says that the concept is indefinite, something which in our language is otherwise, generally expressed with the indefinite pronoun *neki*’ (Stevanović 1964: 1: 319)<sup>269</sup> – Grammar of Serbo-Croatian.

‘Serving as indefinite pronouns are: - the numeral *jedan* in the meaning of ‘neki’ [some] and ‘netko’ [somebody] (Barić et al. 1995: 208)<sup>270</sup> - Grammar of Croatian

‘The numeral *jedan* is often used in the meaning of ‘neki’ [some]’ (Piper and Klajn 2013: 94)<sup>271</sup> - Grammar of Serbian

The comparison of the use and semantics of *jedan* and *neki* (above) and the fact that both items have been identified as being in increased usage as “surrogates for articles” ((Radovanović 2009: 210) is what led to me to investigate and inspect the usage of these items in Serbian and Croatian usage.

In the above quotes from normative grammars we find only a vague comparison to indefinite articles, probably because the grammars are intended for native speakers of article-less languages. However, in descriptions of Serbian and/or Croatian grammar intended for foreign learners of Serbian and Croatian, the analogy with the indefinite article is more explicit:

---

<sup>269</sup> Original: „Број *jedan* се врло често у нашем језику употребљава – не да се њим означи број, него је то онда више као нека врста неодређеног члана. [...] Употребом речи *jedan* у ствари се казује неодређеност појма, оно углавном што се у нашем језику иначе означава неодређеном заменицом *неки* “

<sup>270</sup> Original: “5. Službu neodređene zamjenice imaju: - broj *jedan* u značenju ‘neki’ ili ‘netko’

<sup>271</sup> Original: “Број *jedan* често се употребљава у значењу ‘неки’ “

In a lexicon of grammatical terminology in Serbo-Croatian, written in Danish, it says:

‘...članska UPOTREBA [article usage]; the latter denomination refers to the use of the numeral *one* and pronouns such as *neki, taj, ovaj, onaj* etc. in a way which by and large corresponds to (or is comparable to) the usage of articles in an article-language’ (Barjamović-Orahovski 2006: 1: 44)<sup>272</sup>

In a grammar of Bosnian, Croatian, Serbian, written in English, it is noted that:

“...the meaning of *jedan* is close to that of the English indefinite article (a, an).” (Alexander 2006: 44)

In a grammar of Serbo-Croatian for foreigners, written in Serbo-Croatian, it says:

‘However, the determiner *jedan* may also be used in the function of an article. In Serbo-Croatian the article as a word category, does not exist. [...] If one wishes to emphasise the meaning ‘unknown’, the determiner *jedan* is used.’ (Mrazović and Vukadinović 1990: 243)<sup>273</sup>

In a later edition of this grammar the meaning of ‘new information’ is added to the meaning ‘unknown’ as well as the point that under such circumstances *jedan* is interchangeable with *neki* (Mrazović 2009: 288).

It is curious that, what I see as a necessary distinction between the meaning of the article-like *jedan* and *neki*, has not entered any of the quoted grammars, at least not in an explicit way. Instead, the indefinite pronoun *neki* is frequently used as a means to explain the meaning of the article-like use of *jedan* as is done by Mrazović (2009), Stevanović (1964), Barić et al. (1995), Piper and Klajn (2013).

However, the distinction between the meaning of article-like *jedan* and *neki* was pointed out by Milka Ivić already in 1971 in her article

---

<sup>272</sup> Original: “...članska UPOTREBA; den sidstnævnte betegnelse henviser til brugen af talordet *jedan* og pronominer som *neki, taj, ovaj, onaj* m.fl. på en måde, der i det store hele svarer til (eller lader sig sammenligne med) brugen af artikler i et artikelsprog.”

<sup>273</sup> Original: “Međutim, determinativ *jedan* može da se upotrebi i u funkciji člana. U srpskohrvatskom jeziku ne postoji član kao posebna vrsta reči.[...]. Ako se značenje ‘nepoznat’ želi posebno istaći, koristi se determinativ *jedan*.”



“Leksema *jedan* i problem neodređenog člana”<sup>274</sup>. She concludes that *jedan* and *neki* are very rarely synonymous and, in my paraphrase, that the main difference between the article-like usage of *jedan* and *neki* lies within the semantic fields of SPECIFICITY and REFERENTIALITY. This difference is shown implicitly in Alexander (2006) through her translations of *jedan* into ‘a particular’ and *neki* into ‘some [= a]’ in the following sentences.

- 1) *Tražim jednu osobu* – I’m looking for a [particular] person
  - 2) *Neka devojka te čeka* – Some [= a] girl is waiting for you
- (Alexander 2006: 44)

So, here the difference in semantic content lies in whether the person or the girl is regarded, by the speaker, as a specific girl or a specific person or not. The person (*osobu*) determined by *jedan* is a specific person, whereas the girl (*djevojka*) is determined as any one girl. As shown by Alexander’s translation into English, in both instances the determiner may be seen as the equivalent of an English indefinite article.

In his grammar of Bosnian/Croatian/Serbian, written in Norwegian (Mønnesland 2002) the author does not describe the article-like usage of *jedan*. Mønnesland does, however, use the term ‘indefinite non-specific’<sup>275</sup> in his description of the pronoun *neki* and exemplifies its use with a sentence, in which *neki* is translated into the Norwegian by the indefinite article ‘en’:

- 3) *Došao je neki mlad čovjek* – ‘En ung mann er kommet’  
           came      *neki* young man    - a young man has come

In both grammars (Alexander 2006 and Mønnesland 2002), the use of *neki* is explained and illustrated as the equivalent to an indefinite article in an article-language. Furthermore, both highlight that the non-specificity of the indefinite referent (young man) is encoded in *neki*.

So, we may conclude that manuals of Croatian and/or Serbian written in article languages, which contain contrastive comparisons between the studied language and the language of instruction (as e.g. Danish, Norwegian and English), regularly compare the content and function of *neki*

---

<sup>274</sup> ‘The lexeme *jedan* and the problem of the indefinite article’

<sup>275</sup> Original: “Ubestemt ikke-spesifikt” (Mønnesland 2002: 106)

and *jedan* with the content and function of indefinite articles both explicitly and implicitly (cf. quotes above).

Before we take a look at some of the research of the article-like *neki* and *jedan*, and subsequently the empirical data it is necessary to provide a more exact definition of linguistic (in)definiteness. And, in order to determine the level of “article-like indefiniteness” of *neki* and *jedan* and whether *neki* and *jedan* should be termed ‘indefinite articles’ (in)definiteness encoded in articles must also be defined. The following definition is derived from descriptions of article-languages and exemplifies indefinite usage by showing how utterances involving indefinite articles in English and/or Danish may be rendered in Serbian and Croatian.

#### 4.6.1 Indefiniteness in articles

In languages like Danish and English, definiteness is explicit and obligatorily marked in NPs with articles as in [*a* horse]<sub>NP<sub>indef</sub></sub> and [*the* horse]<sub>NP<sub>def</sub></sub>. In Serbian and Croatian definiteness is not obligatorily marked.<sup>276</sup>

In my native language Danish, the distinction between definiteness and indefiniteness is defined as one of knowledge of and/or hearer’s<sup>277</sup> ability to identify a referent, i.e.:

<u>definite</u>	<u>indefinite</u>
+KNOWN /	-KNOWN
+IDENTIFIABLE	-IDENTIFIABLE

‘By using a *definite* the speaker signals that he believes that the hearer either knows or is able to identify the referred to object’ [...]

‘By using an *indefinite* the hearer is not expected to know or be able to identify the referred to object’ (Hansen 1997: 28)<sup>278</sup>

---

<sup>276</sup> The lack of obligatory marking of (in)definiteness is, according to some, ample proof that the grammatical category definiteness is not present in articles at all. (M. Ivić 1971: 119)

<sup>277</sup> The English terms ‘speaker’ and ‘hearer’ cover the participants in any speech act, written or oral.

<sup>278</sup> Original: “Med bestemt form markerer afsenderen at han mener at modtageren kender eller kan identificere den genstand der tales om. [...] Med ubestemt form markeres det at modtageren ikke forventes at kende eller at kunne identificere genstanden”

Christopher Lyons arrives at a similar conclusion in his monograph *Definiteness* (Lyons 1999) which “investigates definiteness both from a comparative and a theoretical point of view, showing how languages express definiteness and what definiteness is”. He characterises definiteness as signalling the known and the identifiable and remarks that one of the main differences between definite and indefinite NPs lies in the different levels of shared awareness of the referent of that NP. When an indefinite NP is used “the speaker may be aware of what is being referred to and the hearer probably not, with a definite noun phrase this awareness is signalled as being shared by both participants.” (Lyons 1999: 2–3).

The point, that the speaker *may* or *may not* be aware, or rather: signal awareness of what is being referred to is key to understanding the differing indefinite article-like usages of *neki* and *jedan*.

So, the choice to mark an NP as indefinite in Croatian or Serbian comes from the speaker’s wish to show that s/he is aware that the hearer does not know or is unable to identify the referent of the NP. After having chosen to mark an NP as an indefinite, the speaker then has to choose between the indefinite markers (determiners). This choice will be reliant upon the speaker’s level of (signalled) awareness of the referent of the NP, and not upon the expected knowledge of the hearer.

In this respect the Serbian and Croatian article-like *neki* and *jedan* differ from indefinite articles in article-languages such as Danish or English. In these languages the indefinite article does not signal whether the referent of an NP is known to the speaker as is evident in the following sentences:

“I’m looking for a record.” (Lyons 1999: 35)

Danish equivalent:

‘Jeg leder efter en plade.’

The speaker may be looking for a specific record or any record (i.e. a non-specific record).

In Croatian or Serbian a speaker can either omit a determiner (4a), or use a determiner (4b, 4c) which then signals the specificity of the record in the eyes of the speaker:

4a) *Tražim ploču.*

I search record

'I am looking for the record.' (+DEFINITE, +SPECIFIC)

'I am looking for a record.' (-DEFINITE, +/- SPECIFIC)

4b) *Tražim jednu ploču.*

I search JEDAN record

'I am looking for a record.' (-DEFINITE, + SPECIFIC)

4c) *Tražim neku ploču.*

I search NEKI record

'I am looking for a record.' (-DEFINITE, -SPECIFIC)

4d) *Tražim tu ploču.*

I search TAJ record

'I am looking for the record.' (+DEFINITE, +SPECIFIC)

'I am looking for that record.' (+DEIXIC, +DEFINITE, +SPECIFIC)

In sentence (4a) no kind of definiteness or specificity is marked so the record may be a record which is identifiable to the hearer, i.e. 'the record'- DEFINITE, or unidentifiable to the hearer, i.e. 'a record' - INDEFINITE.  
279 So, whether the speaker is referring to a specific record or any record in the 'indefinite' reading of *ploču* in (4a) is not marked.

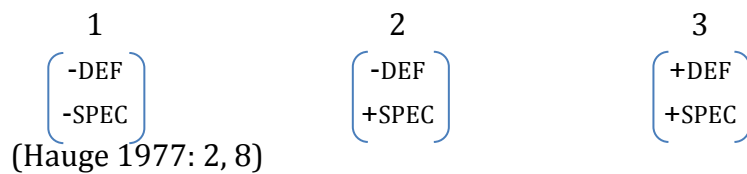
In sentence (4b), however, the record in question is *a specific record* which the hearer is not expected to be able to identify. In sentence (4c) the speaker marks the record as unidentifiable and *non-specific*. It is worth noting that the article-less (4a) is not the only sentence where *ploču* may be interpreted as a definite 'record'. Katunar et al. inform us that even the deixic (demonstrative) pronoun *taj* may be used as the equivalent of the definite article which would result in the two possible readings of sentence (4d) (Katunar et al 2013: 31). So, the means to mark a NP as either definite or indefinite are present and in use in Serbian and Croatian and this *may* lead to a full grammaticalisation of definiteness in the shared linguistic system of Croatian and Serbian.

---

279 The record will, in the mind of the speaker, certainly be either a specific, definite, specific, indefinite or unspecific, indefinite record. But, the level of speaker's awareness of the referent (the record) will be signalled by other textual or contextual means.

Heine and Kuteva (2006) who state that “articles are spreading throughout Europe.” also use specificity as a way to delineate one kind of indefinite marking from another, and claim that the “specific, indefinite marking” of an NP with the numeral ‘one’ is one of the functions ‘one’ takes on in its grammaticalisation which may lead it to a stage where it can be considered a marker of the “nonspecific, indefinite”. (Heine and Kuteva 2006: 97, 104–5). Heine and Kuteva consider the non-specific, indefinite marking as a more grammaticalised indefinite usage than the specific one. They do, however, only consider the numeral ‘one’ and do not include any other possible markers of indefiniteness in their description of “The rise of articles” in several languages, including Serbian and Croatian.

Breu (2012) also focuses on specificity in his description of a contact-induced grammaticalisation of the numeral ‘one’ in Sorbian and Molise Slavic but uses the terms *referential*, *non-referential* and *generic*. And, Hauge (1977) uses *specificity* and *referentiality* as descriptors of the content of Bulgarian *една* (‘one/a’) when it is “a potential indefinite article” and proposes to divide NPs into three groups according to the features definiteness [+/-DEF] and [+/-SPEC], like so:



These terms are not without problems. Often, non-specific (-SPEC) is considered synonymous with generic (+GEN), which is an inadequate description. Consider:

- 5a) George became a teacher (-DEF, -SPEC, +GEN)  
 5b) George asked a teacher (-DEF, -/+SPEC, -GEN)

In (5a) ‘a teacher’ does not refer to a person, but a general idea (a profession) whereas in (5b) the referent of ‘a teacher’ is an existing subject, of which is not marked either known (+SPEC) or unknown (-SPEC) to the speaker.<sup>280</sup>

---

<sup>280</sup> In Danish the article is omitted when an NP is generic: *Han blev lærer*  
 ‘he became teacher’

It is also problematic to state that all definite NPs are also specific NPs as is the import of Hauge's group 3 above.

5c) George asked the teacher (+DEF, +SPEC, -GEN)

5d) George went to the doctor (+DEF, +/-SPEC, -GEN)

5e) The lion is a dangerous animal (+DEF, -SPEC, +GEN)

As Lyons (1999: 7–12) points out, the definite article is also used to determine NPs that do not refer to a specific NP as would be the case in one of the possible readings of (5d), that is, in situations where the doctor is not a doctor who is known to the speaker, but still an individual doctor, and therefore not a generic reference as in (5e). As *jedan* and *neki* are not used to mark definiteness, I will not discuss the semantic scope of definiteness [+DEF] any further, but focus on indefiniteness and refer to some of the existing research on grammatical pattern replication concerning indefinite articles.

#### **4.6.2 Research of article-like usage of *jedan* and *neki***

As mentioned above, Heine and Kuteva (2006, 2010) have described the indefinite article-like use of the numeral 'one' as a process of grammaticalisation, and as part of „The Rise of Articles“<sup>281</sup>. Furthermore, they argue that the grammaticalisation of articles in Slavic languages is contact-induced (Heine and Kuteva 2010: 91). Finally, they believe that „the following [five] stages [...] mark the gradual pragmatic and semantic evolution of many indefinite articles:

1. An item serves as a nominal modifier denoting the numerical value 'one' (numeral).
2. The item introduces a new participant presumed to be unknown to the hearer and this participant is then taken up as definite in subsequent discourse (presentative marker).
3. The item presents a participant known to the speaker but presumed to be unknown to the hearer, irrespective of whether or not the participant is expected to come up as a major discourse participant (specific indefinite marker).
4. The item presents a participant whose referential identity neither the hearer nor the speaker knows (non-specific indefinite marker).

---

<sup>281</sup> The heading of chapter 3 in Heine and Kuteva (2006)

5. The item can be expected to occur in all contexts and on all types of nouns except for a few contexts involving, for instance, definiteness marking, proper nouns, predicative clauses, etc. (generalized indefinite article).“ (Ibid: 91-92)

The four stages (2-5) are all functions of indefinite articles. The stages may be viewed as an evolutionary scale, which is also an implicational scale, i.e. if the numeral 'one' is used as a specific indefinite marker (stage 3) it will also be used as a presentative marker (stage 2) (Ibid: 92). Their own research of colloquial Serbian and Croatian, presented in Heine and Kuteva (2006) led them to place colloquial Serbian and Croatian partly on stage 2 and partly on stage 3 (Heine and Kuteva 2006: 125).

The grammaticalisation of the numeral 'one' into the indefinite article, or the use of 'one' as an indefinite article is a well-known phenomenon across article-languages. Consider, for instance, Turkish: *bir*, French: *un*, German: *ein* and the Danish/Norwegian/Swedish: *en*. Therefore, it is likely that *jedan* would also become an indefinite article.

However, it seems that in Serbian and Croatian the semantic field of indefinite articles (i.e. marking an NP as unknown or unidentifiable to the hearer) is covered by at least one more lemma: *neki*. The primary difference between *jedan* and *neki* when used as indefinite articles lies, as shown above, in the specificity of the referent.

This claim is supported by researchers who have investigated the uses of *jedan* and *neki* but is not, as we saw, evident in normative works of Serbian and Croatian. A thorough account is given by Milka Ivić (1971) who states, as opposed to the normative works, that *neki* and *jedan* are very rarely synonymous (M. Ivić 1971: 118). On the contrary, *jedan* is used when something is marked for its 'individuality' as opposed to *neki* which is specialized for situations with a 'general lack of definiteness' (Ibid: 112). Furthermore, *jedan* signals 'individuation' and 'particularisation' whereas *neki* is used to 'insist on the factual lack of knowledge' and to 'underline that the referent is unidentified'. Finally, both items may be used 'to introduce new information' and it is on these occasions that they are almost synonymous (Ibid: 107, 108, 110, 117 and 111).

Applying the terminology introduced above, what Ivić tells us is that *jedan* marks the specific and *neki* marks not only the non-specific but also an emphasised arbitrariness, which could be illustrated as a semantic feature [+UNSPEC] thereby indicating the emphasis on the lack of

specificity as opposed the mere absence of specificity [-SPEC]. I will, however, continue to use [-SPEC] in order to avoid a further complication of the applied terminology. Ivić also informs us that *jedan*, according to the rules of standard Serbo-Croatian, may not/cannot be used to determine the NP which refers to something which is known to both participants in the speech act, i.e. given information. Nor may it be used to determine a NP which refers to something general, i.e. a generic reference (M. Ivić 1971: 104). In other words: *jedan* cannot determine a 'definite' referent, viz. NP<sub>[+SPEC, +DEF]</sub>, nor can it determine a generic referent, viz. a NP<sub>[-SPEC, +GEN]</sub>. 282

Katunar et al. (2013) who examine specificity markers in Croatian, conclude: "Thus it seems that *jedan* 'a; one' is used predominantly as a specific indefinite marker [...] [And] the indefinite pronoun *neki* 'some' shows a tendency towards indefinite non-specific readings." (Katunar et al. 2013: 37). However, they also highlight the fact that *neki* may be used to mark a specific referent, but only when this specificity is contextual, i.e. marked not by *neki* itself but for instance by a relative clause as in the following sentence:

(6) *Neki čovjek, koji se spremao na dugo putovanje, razmišljao je kako da smanji trošak*

'A man, who was preparing for a long journey, was thinking about ways to minimise costs.'

(Ibid: 36)

The relative clause (underlined) is what makes the referent of *Neki čovjek* ('A man') a specific referent. It is in exactly such situations that M. Ivić finds *neki* and *jedan* nearly synonymous, i.e. when they are used to introduce a new referent which is made specific through contextual means. (M. Ivić 1971: 111) Furthermore, it is this function of an indefinite article, that Heine and Kuteva label as a "presentative marker" and it is presumably the same use that Mrazović (2009) refers to when she says that *jedan* and *neki* are interchangeable when used to introduce new information (Mrazović 2009: 288).

---

<sup>282</sup> The term generic reference is used here as covering both reference to all referents of a given class and predicative use as in: 'Lav je životinja.' (The lion is an animal, i.e. all lions are animals) and 'On je postao učitelj.' (He became a teacher, i.e. he is now a member of the class of referents who are teachers)



The possibility of a ‘contextual specificity’, however, does not mean that *neki* itself encodes specificity of the NP which it determines. Therefore, it stands that *jedan* is a specificity marker: [+SPEC] whereas *neki* marks non-specificity: [-SPEC]. By combining Heine and Kuteva’s definition of ‘one’ as a presentative marker with Katunar et al.’s description of contextual specificity we may speak of a *jedan* and *neki* used as presentative markers regardless of whether they also mark specificity, i.e.: [+PRES, +/-SPEC]. So, the existing research points at the following possible article-like usages of *jedan* and *neki*.

Figure 4.4 - Indefinite properties of ‘jedan’ and ‘neki’ in NPs

	<i>jedan</i>	<i>neki</i>
presentation of participant [+PRES]	+	+
specific reference [+SPEC]	+	-
non-specific reference [-SPEC]	-	+
generic reference [+GEN]	-	-

I have found no comparative research of Croatian and Serbian concerning *jedan* and *neki*. Neither does any of the consulted sources on differences between Serbian and Croatian nor the grammars which describe both Croatian and Serbian mention any difference between the two languages in this respect<sup>283</sup>. To confirm this apparent sameness I have searched the corpora for similar examples.

Let us now proceed to see whether the usage norm as represented in the SrpKor and HNK will confirm this claim as well as the claims presented in Figure 4.4.

#### 4.6.3 Empirical data

In order to limit the quantity of results and at the same time narrow the focus of the searches I have chosen to search only for NPs consisting of two items, i.e. a noun preceded by either *neki* or *jedan*, thus avoiding NPs with attributive modifiers which would make it more difficult to determine whether the specificity or unspecificity of an NP is encoded only in the determiner (*jedan* or *neki*). Furthermore, I have narrowed my search to NPs in the singular in order to gain comparability with article languages such as English or Danish which do not have indefinite articles in the plural. Finally, I have refrained from searching for NPs with proper

---

<sup>283</sup> Among others: (B. Brborić 2011; Bekavac et al. 2008; Langston and Peti-Stantić 2014; Mønnesland 2002; Alexander 2006; Tošović 2010)

nouns. Proper nouns are definite per se in as much as they refer to an identifiable referent. So when *jedan* and *neki* are used as determiners of proper nouns, they take on the meaning: “someone like” and “someone called”, respectively.

- 1) “*kad govorimo o filozofiji jednog Dekarta*” (M. Ivić 1971: 110)  
‘when speaking about the philosophy of someone like Descartes’
- 2) “*tip ima talent jednoga Šopena*” (M. Ivić 1971: 115)  
‘the guy has a chopinesque talent’
- 3) “*na vrata uđe i priđe k nama neki Sima Martin*” (Ibid: 111)  
‘someone called Sima Martin entered and came towards us’

I do not consider it necessary to search for more than a few different common nouns in the singular in order to ascertain what kind of indefinite article-like functions are part of the usage norm of Serbian and Croatian as documented in SrpKor and HNK. I have therefore searched the two corpora for:

<i>jedan prijatelj</i>	<i>neki prijatelj</i>	‘friend’
<i>jedna djevojka</i>	<i>neka djevojka</i>	‘girl’
<i>jedan stol</i>	<i>neki stol</i>	‘table’
<i>jedna zgrada</i>	<i>neka zgrada</i>	‘building’

I have searched for countable nouns with concrete referents which I consider to be frequently used: Two which have human referents and two which have inanimate referents.

#### *Search criteria*

All queries conducted in both corpora were compiled in concordances which I inspected in order to determine whether *jedan* and *neki* are used in the usage norms of Croatian and Serbian as indefinite articles by marking a noun as:

- a) a participant presumably unknown to the hearer which is introduced as such only to be contextually specified, i.e. taken up again immediately thereafter (presentative) [+PRES].
- b) a participant known to the speaker but presumed to be unknown to the hearer (specific) [+SPEC]

c) a participant whose referential identity neither the hearer nor the speaker knows (non-specific) [-SPEC]

d) a participant which has a generic reference [-SPEC, +GEN]

The compiled concordances supplied me with 1137 results in the SrpKor and 465 in HNK. A number of the concordances were excluded from further analysis. Some were excluded because they were homographs of the NPs in question, e. g. the form *neke djevojke* may be *neka djevojka* in the nominative plural ('girls') as well as in the genitive singular ('a girl'). Another homograph is the adverb *jednom* ('once') which is orthographically identical to the *jedan* in the instrumental, feminine, singular as in (1) and in the locative and dative, masculine/neuter singular, as in (2):

- 1) "*Česlav Miloš je [...] zapodenuo sasvim neobavezni razgovor sa jednom devojkom koja je, takođe, nekoga čekala*" (SrpKor)  
(Česlav Miloš initiated a completely non-committing conversation; also waiting for somebody')
- 2) "*[Irena] poče premetati neke knjige na jednom stolu*." (HNK)  
(Irena began to rearrange some books on a table).

Further, *neka* (nominative, singular, feminine of *neki*) has a homograph - a particle carrying the imperative meaning "Let!" as in "*Neka djevojka ide*" 'Let the girl go!' as opposed to the indefinite article-like or indefinite pronominal reading: 'A/Some girl is going/coming'. Finally the wordform *nekome* (locative, singular, masculine of *neki*) has a homograph which is the dative of the nominal pronoun *neko* ('somebody').<sup>284</sup>

Other concordances were excluded from further analysis not because the examples did not contain a form of *jedan* or *neki* in the singular, but because they were not examples of the article-like usage of *jedan* and *neki*. Instead they were examples of numerical usage of *jedan* ('one') or the pronominal usage of *neki* ('some') as in (3) and (4):

- 3) "*Dva kreveta, jedan stol i tri stolca*." (HNK)

---

<sup>284</sup> As *neko* is a nominal pronoun, it is never the attributive non-head of an NP as is *neki*. This feature makes the nominal pronoun *neko* easy to detect.

('Two beds, one table and three chairs')

- 4) "Pravi mi se neki prijatelj, a smatra da sam još uvek balavac!"  
(SrpKor)  
(‘He pretends to be some kind of friend to me, but he thinks I’m still a snotty kid!’)

Even with the exclusion of such concordances a large number samples containing indefinite article-like usage of *jedan* and *neki* remain for further analysis. So, having excluded the mentioned concordances the remaining were further inspected and analysed, and this analysis shows that it is amply evident that both *jedan* and *neki* are widely used in a function identical to that of indefinite articles and, furthermore, there seems to be no apparent differences between Croatian and Serbian usage.

*The presentative function*

*jedan* is used as a presentative marker [+PRES], i.e. to introduce a participant as new, unknown information [-DEF], which is immediately taken up again (co-referred to), as described by several researchers, cf. above. And, *neki* may also take on this function, as pointed out by Mrazović (2009). Among the concordances of the eight queried NPs, I found several examples of [JEDAN+N]<sub>NP(+PRES)</sub> and [NEKI+N]<sub>NP(+PRES)</sub> <sup>285</sup>:

- 5) [...]Gaudi posvećuje mu čak čitavu jednu zgradu, čuvenu Casa Batlló"  
(SrpKor)  
(‘...Gaudi even dedicates an entire building to him<sup>286</sup>, the famous Casa Batlló’)
- 6) I u tom[...] gradu ima jedna djevojka za koju se misli da je bjelija od snijega (HNK)  
(‘And in that city there is a girl of whom it is thought that she is whiter than snow’)

---

<sup>285</sup> The queried NPs are underlined as well as the items in the examples which constitute the further mentioning of the same participant in the discourse.

<sup>286</sup> St. George

- 7) [...] *zašto kad razmišljamo o nekom prijatelju s kojim smo izgubili svaku vezu, uspevamo da ga iznenada sretnemo* (SrpKor)  
(‘why is it that, when we think of a friend, with whom we have lost all contact, we succeed in suddenly happening upon him’)
- 8) *Kupio ju je u Splitu za 1000 kuna od nekog prijatelja s kojim je bio u zatvoru* (HNK)  
(‘He bought it in Split for a thousand *kuna* from a friend with whom he had been in prison’)

In all four examples the referent of the noun in the NP is a specific referent, i.e. a specific building (5), a specific girl (6) and a specific friend (7, 8). The difference between the seemingly synonymous usage of *jedan* and *neki* in these examples lies in the fact that the building (4) and the girl (5) are marked as specific by their determiner – *jedan*, and then mentioned immediately thereafter whereas the friend in (7) and (8) only obtain specificity through the subsequent information in the sentences, i.e. ‘the friend with whom we have lost all contact’ and ‘the friend with whom he had been imprisoned’. So, the specificity and non-specificity marked by the use of *jedan* and *neki*, resp. is present in example 5-8 but the non-specificity of the noun determined by *neki* is annulled by the co-referentiality of *prijatelju* (‘friend’) and *kojim* (‘whom’) and *ga* (‘him’) in (7) and by the co-referentiality of *prijatelja* (‘friend’) and *kojim* (‘whom’) in (8). This introduction of new information followed by a co-reference to the same referent confirms the findings of Heine and Kuteva (2006, 2010) concerning the use of *jedan* as a presentative marker and of Katunar et al (2013) concerning that not only *jedan* but also *neki* is used as an indefinite article, functioning as a “presentative marker”. This does not, in my view, make *neki* and *jedan* synonymous when used to introduce new information, it merely tells us that both determiners may be used for this purpose, i.e. function as presentative markers.

#### *Marking specificity*

Both corpora contain examples of *jedan* used as a specificity marker of the referent of the four queried NPs without any immediate further reference to this discourse participant:

- 9) *Ti su dokumenti[...]nađeni dok su trunuli u jednoj zgradi u Banjoj Luci* (HNK)  
 'Those documents were found rotting in a building in Banja Luka'
- 10) *Društvo se posadi za jedan sto*. (SrpKor)  
 'The friends planted themselves at a table.'
- 11) *Jedamput mu je jedna djevojka rekla* (HNK)  
 'Once a girl told him'
- 12) *Čekala vas je jedna devojka*. (SrpKor)  
A girl was waiting for you

By analysing the context in the two corpora surrounding examples (9)-(12), I could conclude that the function of *jedan* in these examples is non-numerical and, in addition, it marks the participant as indefinite, i.e. presumed unknown to the hearer. It also serves to tell us that the 'building', 'table' and 'girl' in question are, in the mind of the speaker [+SPEC]. In short, all instances of [JEDAN+N]<sub>NP</sub> in (9)-(12) may be described as [JEDAN+N]<sub>NP</sub>(+SPEC, -DEF)

I found no instances in which *neki* marked an NP as [+SPEC] though some occurrences were a little bewildering, as for instance:

- 13) *Onaj koji voli neku djevojku, voli i njezinu kosu i njezine oči i usne i njezine ruke* (HNK)  
 'He who loves a girl, also loves her hair and her eyes and lips and her hands'

What is bewildering is that one is inclined to think that the girl whom he loves must be a specific girl. And, surely she is, but, in (13), as in (7) and (8) the referent of the [NEKI+N]<sub>NP</sub> only becomes marked as specific through the subsequent co-references, i.e. 'her hair', 'her eyes', etc. And even then, the referent may be said to stay non-specific in as much as the girl who is the object of love could be any girl, i.e. not a specific or certain girl. She is, however, an individual girl and not a representative of all girls. So, there is no truly generic reference in (13) although in its English translation, 'a girl' could have a generic reference.

So, as a marker of specific indefiniteness, I only found [JEDAN+N]<sub>NP</sub>'s which supports the existing research of the non-numerical properties of

*jedan*. As opposed to Heine and Kuteva (2006) who regard Serbian and Croatian article-like use of *jedan* primarily as [+PRES] and not fully as [+SPEC], I consider the indefinite specificity [+SPEC] as the primary function of article-like *jedan*

#### *Marking non-specificity*

In the two corpora I found a large number of examples of [NEKI+N]<sub>NP</sub> marking an indefinite and non-specific participant, i.e. a participant whose referential identity neither the hearer nor the speaker knows. (13) above is one such example. Consider also examples (14)-(16):

- 14) *Uvek je tražio neki sto u ćošku, nije voleo da bude zapažen.* (SrpKor)  
'He always looked for a table in the corner, he didn't like to be noticed.
- 15) *Povremeno bi mu i neki prijatelj donio litru domaće rakije.* (HNK)  
'Sometimes a friend would bring him a litre of homemade brandy'
- 16) *Znači, trebalo je samo da nađe neku devojku, pravu devojku od krvi i mesa* (SrpKor)  
'So, he only needed to find a girl, a real flesh-and-blood girl'

The 'table' in (14) is not a specific table, nor does it have to be a specific 'friend' who brings the brandy in (15), and the 'girl' in (16) is certainly not a specific girl. The use of *neki* in (15) shows how important the speaker's perspective is when determining the presence or lack of specificity. The friend in (15) could very well be a certain friend whose identity the speaker knows, but who s/he chooses to mark as a non-specific friend, thereby deliberately obscuring the speaker's ability to identify said friend.

The many examples in the corpora of non-specific usage of *neki* combined with the four nouns, confirm the claim put forward by Katunar et al. (2013) that "the indefinite pronoun *neki* 'some' shows a tendency towards indefinite non-specific readings." Katunar et al. (2013: 37).

However, in the corpora, I also found a few examples of non-specific use of *jedan*:

- 17) *kakve su moralne kvalifikacije potrebne da jedna djevojka bude proglašena "bijelom robinjicom"* (SrpKor)  
'which are the required moral qualifications for a girl to be proclaimed a "white slave" '
- 18) *to je tada za jednu djevojku iz Hrvatske bilo strašno puno* (HNK)  
'back then, that was extremely expensive for a girl from Croatia'
- 19) *Rezervirao je jedan stol dolje u baru. To nek je ništa ne smeta, jer restoran je pun svakakvoga svijeta , a u baru je intimnije i ugodnije.* (HNK)  
'He booked a table down in the bar. That shouldn't bother her, because the restaurant is full of all kinds of people and it's more comfortable and private in the bar.'

In (17) and (18) the 'girl' is a non-specific girl. It could in other words, be any girl. In (19) the 'table' could be interpreted as a specific table, but the wider context in which the concordance was found does not indicate that, nor does it indicate any emphasis on the numerical meaning of *jedan* – 'one'.

This usage of *jedan* does not comply with the descriptions of the possible usages of *jedan* provided by M. Ivić (1971) nor does it fit the descriptions given in the grammars for foreigners. Only the description (but no examples) provided by (Katunar et al. 2013) allows for non-specific indefinite use of *jedan*: "This is not to say *jedan* 'a; one' has to be a specific marker in every instance of its usage, but only that it shows a strong tendency towards specific reading." (Katunar et al. 2013: 37).

In my view, this usage also confirms that Serbian and Croatian possess two linguistic items (*jedan* and *neki*) which may be defined as markers of indefiniteness when the languages are compared to article languages or taught to speakers of article languages, but the essential semantic properties of these items is specificity understood as speaker's knowledge of the referent.

#### *Marking a generic reference*

During the inspection of all the concordances of *neki* and *jedan* as determiners of the four selected nouns in both corpora I found no examples



of *neki* or *jedan* in NPs with a generic reference. In order to avoid a faulty conclusion concerning the possible generic use of [NEKI+N]<sub>NP(+GEN)</sub> and [JEDAN+N]<sub>NP(+GEN)</sub> I conducted additional searches of both corpora querying phrases as "POSTAO+JEDAN+N" and "POSTAO+NEKI+N" both equivalent to ('BECOME a N')<sup>287</sup>.

Since the intention was to discover whether there were any instances of a [NEKI+N]<sub>NP</sub> or a [JEDAN+N]<sub>NP</sub> referring to all the possible referents or to a general idea (exemplified by "5a) George became a teacher" on p. 262), this approach, seemed suitable.

In the SrpKor I got a sample of 113 concordances but in HNK only 14. The uses of *jedan* were all numerical, i.e. translatable with 'one' whereas the uses of *neki* were all examples of the unspecific pronominal usage, translatable as 'some kind of', 'some sort of', etc., e.g.:

- 20) *Posle izvesnog vremena je pristao, ne prestajući da me gleda kao da sam mu odjednom postao neka pretnja* (SrpKor)

'After a certain amount of time he accepted but he didn't stop looking at me as if I had suddenly become some sort of threat to him'

- 21) *"Sagradih ovaj grad i dovedoh u nj narode iz cijeloga moga kraljevstva , razne narode , raznih jezika, da bi u ovome gradu postali jedan narod, narod jednog jezika"* (HNK)

"I built this city and into it I brought nations from all over my kingdom, different nations of different tongues, so that they in this city could become one nation, a nation of one tongue."<sup>288</sup>

So, after having carefully inspected all the concordances in the initial queries, exemplified here by (1)-(19) and having carried out additional searches I have come to the conclusion that in Croatian and Serbian usage norm in the publicistic functional style, *jedan* and *neki* are not used as indefinite articles when the referent of an NP is generic. However,

---

<sup>287</sup> Cf. Appendix D for detailed description of POSTAO

<sup>288</sup> Though published in the daily newspaper *Vjesnik*, this quote is not typical for the publicistic functional style. It is a translation of the inscription on a tablet found in an Ancient Assyrian city.

the remaining functional features of an indefinite article are covered by *jedan* and *neki*.

### ***Summing up jedan and neki***

When a speaker of either Serbian or Croatian wants to mark a participant as indefinite (-DEF) (i.e. unidentifiable to the hearer) and either known (+SPEC) or unknown (-SPEC) to the speaker s/he may use *jedan* to encode (-DEF, +SPEC) and *neki* to encode (-DEF, -SPEC). On this basis we may conclude that both *jedan* and *neki* have additional functions beside their basic numerical and pronominal function. These additional functions *jointly* cover most of the functions ascribed to indefinite articles in article-languages. So, *jedan* and *neki* do function as equivalents to indefinite articles in article-languages. This use of *jedan* and *neki* is relatively new. The usage of *jedan* in this presumed contact-induced article-like function was noticed over a hundred years ago. And the similar use of *neki* has been recognised and compared to the use of indefinite articles in the Germanic languages. Furthermore, in both cases, their article-like use is not a differential marker between Croatian and Serbian as documented in HNK and SrpKor, although it has been presumed that the article-like use of *jedan* was a particularly Croatian trait stemming from the Kajkavian and Čakavian dialects spoken in Croatia (cf. p. 141).

### ***Grammaticalisation of jedan and neki***

The fact that speakers of Croatian and Serbian use *jedan* and *neki* in the manners described and exemplified above is evidence of an on-going grammaticalisation where a grammatical pattern (marking indefiniteness) has risen from being a *minor* to a *major grammatical use pattern* possibly due to a contact-induced *replication of grammatical use patterns*, just as envisaged by Heine and Kuteva (2005):

#### ***The rise of a major use pattern in contact-induced replication***

- a. An existing use pattern is used more frequently
  - b. It is used in new contexts
  - c. It may become associated with a new grammatical function
- (Heine and Kuteva 2005: 45)

As this study shows, the possible grammatical pattern replication of indefiniteness should not be investigated by only looking at one potential candidate to fill out the role of indefinite article, but one should rather keep in mind that the grammatical category of indefiniteness may be expressed with different means within the same language. In other words,

the fact that *jedan* has not grammaticalised into an indefinite article which can be used to determine specific as well as non-specific NPs does not mean that indefinite non-specificity is not marked in Serbian and Croatian at all. This function of the indefinite article has been replicated on to *neki* instead. So, by using the determiners *jedan* and *neki*, speakers of Serbian and Croatian are able to convey three out of four functions typical for indefinite articles in article-languages as shown in Figure 4.4 (introduced on p. 266):

Figure 4.4 - Indefinite properties of *jedan* and *neki* in NPs

	<i>jedan</i>	<i>neki</i>
presentation of participant [+PRES]	+	+
specific reference [+SPEC]	+	-
non-specific reference [-SPEC]	-	+
generic reference [+GEN]	-	-

Having claimed this, we must also acknowledge that while most features of indefinite articles have been replicated, the grammatical category ‘definiteness’ has not been replicated neither in Croatian nor in Serbian. Definiteness in article-languages entails obligatory marking of definiteness, but marking of (in)definiteness while being possible, is *not* obligatory in Croatian and Serbian. Applying the well-known words of Roman Jakobson<sup>289</sup>: Indefiniteness is something that *may* be conveyed but not something that *must* be conveyed in Serbian and Croatian.

Focusing on the lemmas *jedan* and *neki* we may, however, talk about grammaticalisation where both lemmas have passed through at least two out of the four<sup>290</sup> stages of grammaticalisation as lined out in chapter 3, p.120 and repeated below:

---

<sup>289</sup> “Languages differ essentially in what they *must* convey and not in what they *may* convey” (Jakobson 1959: 236)

<sup>290</sup> Parameter d) *erosion* is, unsurprisingly, not attested in corpora consisting of *written language* (cf. the classification of HNK and SrpKor on p. 52)

*Parameters of grammaticalisation:*

- a) extension, i.e. the rise of novel grammatical meanings when linguistic expressions are extended to new contexts (context-induced reinterpretation)
- b) desemanticizing (or “semantic bleaching”), i.e. loss (or generalization) in meaning content
- c) decategorialization, i.e. loss in morphosyntactic properties characteristic of lexical or other less grammaticalised forms, and
- d) erosion (or “phonetic reduction”), i.e. loss in phonetic substance (Heine and Kuteva 2005: 15)

Thus *jedan* has (a) taken on new grammatical meaning, i.e. [–DEF, +SPEC]. When used as a marker of [–DEF, +SPEC] the quantifier meaning of *jedan* has been (b) generalised (“bleached” or desemantized) from ONE to A PARTICULAR. There is no loss in (c) morphological properties, i.e. the determiner *jedan* is declined in the same way as the numeral *jedan*. There is, however loss in (c) syntactic properties, i.e. *jedan* when used to mark the indefiniteness of a noun, does not regularly stand alone, i.e. without an explicit noun as the head of the NP, whereas the numeral *jedan* may very well stand alone, as in (22) and (23):

- 22) *I to ne samo zbog napada, jednog u nizu* (SrpKor)  
‘And that is not just because of the attack, one in a row (of many).’

- 23) *u Španiji Ministarstvo brine i o centrima za olimpijske pripreme (ima ih tri) a mi nemamo ni jedan.* (SrpKor)  
‘in Spain the ministry even provides centres for Olympic preparations (there are three of them) and we do not have even one.’

The grammaticalisation of *neki* is more difficult to pinpoint as the *indefinite pronoun* that underlines the unspecificity [+UNSPEC] of the NP which it determines is functionally very close to being a marker of indefiniteness and non-specificity [–DEF, –SPEC]. The subtle difference between the unspecificity (emphasised lack of specificity) of the pronoun and non-specificity (general lack of specificity) of the article-like *neki* could be seen as the fulfilment of parameter (a) extension and (b) the generalisation and semantic bleaching of *neki*.

We may also illustrate the grammaticalisation of *jedan* and *neki* using the grammaticalisation cline, introduced as *Cline b* on p. 120

*Cline b*

no function > lower degree of function > higher degree of function

	<i>neki</i> <sub>PRON</sub>	>	<i>neki</i>	} ART, indef
<i>jedan</i> <sub>NUM</sub>		>	<i>jedan</i>	

The quantifier *jedan* is, in my view, a more lexical item, i.e. has more content, namely the meaning ONE, than does the indefinite pronoun *neki* whose meaning can only be explained with formal and functional descriptions as INDEFINITE PRONOUN and UNSPECIFICITY AND INDEFINITENESS OF THE DETERMINED NOUN PHRASE. The curly bracket indicates that *neki* and *jedan* have been grammaticalised to serve functionally (but not obligatorily) in Croatian as Serbian as the equivalent of indefinite articles in article-languages.

## 5. Recapitulation and Conclusions

The preceding chapters contain discussions of theory and empirical evidence related to the question of contact-induced linguistic innovations with special reference to Serbian and Croatian. The form and linguistic implications of these innovations are described and discussed, their acceptance and implementation within the usage norm of the languages analysed. The overarching research question of the thesis is:

*Do the linguistic effects of globalisation enhance or diminish the differences between Serbian and Croatian?*

In this concluding chapter I will present the most important insights which I reached in my attempt to answer this question.

### 5.1 Language planners and language users

**Language planning** played an important role in the unification of Serbian and Croatian into one standard language, the Serbo-Croatian language. The motivation behind the initial unification, from the mid-19<sup>th</sup> century and until the interwar period, was political. The **standardisation** of the common norm of Serbo-Croatian was inspired by aspirations to political emancipation from the Austro-Hungarian Empire under whose rule the Serbs and Croats of today's Croatia, Northern Serbia (Vojvodina) and Bosnia-Herzegovina were living at the time.

The diversification and subsequent partition of Serbo-Croatian in the last decades of the 20<sup>th</sup> century was also politically motivated. This time, Croats (and later Bosniacs) aspired to emancipate themselves from Yugoslav rule. The dominating language planning efforts leading up to, surrounding and following the pivotal political changes of the two world wars and the wars in the 1990s were aimed at achieving the same (opposing) goals as those of the political leaders of the Serbs and Croats, namely unification or separation. During the break-up of Yugoslavia the **status planning** of Croatian and Serbian led to their inauguration as official languages in Croatia (1990) and Serbia (1992)<sup>291</sup>.

During and following the wars in 1990s the **corpus planning** of especially Croatian, in the form of normative works and recommendations

---

<sup>291</sup> And later in Bosnia-Herzegovina (Bosnian, Croatian and Serbian) and Montenegro (Serbian and Montenegrin)

from **linguistic authorities**, was permeated by a socio-political tendency towards 'recroatisising' Croatian, i.e. making standard Croatian more Croat than it had been during its coexistence with Serbian. The Croatisation meant the demotion of all that was labelled Serbian or otherwise foreign in language along with the simultaneous promotion of everything considered to be Croat. A long puristic tradition within Croatian language planning was hereby revived and intensified.

The Serbian situation was very different. The new official status of Serbian, albeit also socio-politically motivated, did not lead to any significant changes in corpus planning. The enhanced differences between recommended Serbian and Croatian language usage of the 1990s were among Serbs rejected as artificial creations made up by Croats. Besides protesting against the new naming of Serbo-Croatian by Croats, Bosniacs and Montenegrins, the language planners of Serbian primarily concerned themselves with questions of status of the Cyrillic script and the Anglification of the Serbian lexis.

The changes in status from varieties<sup>292</sup> of one standard language to two standard languages and the changes in the planned corpus all concern the **prescribed norm of the standard language**. Furthermore, many of the changes of the Croatian standard language were introduced because of their **symbolic function** and did not strengthen the **communicative function**. Instead, this primary function of a language, was weakened, as made evident in numerous cases of mumbling politicians afraid that they would accidentally use a wrong word (i.e. one labelled as Serbian).

Whether the presented changes in the standard language are also implemented by the **language users** is an entirely different matter. Generally, one of the key agents of spreading linguistic innovations (recommended by linguistic authorities or not) is the mass media. So, in order to determine what sort of impact the language planning and language policy of linguistic authorities concerning linguistic innovations have on language usage and, in order to discover to what degree the efforts of planners and policy makers to promote or demote the use of such innovations is successful, I consulted two electronic text corpora representing Serbian and Croatian **written usage norm** in the **publicistic functional style**, i.e. the language of the media.

---

<sup>292</sup> Recognised among Serbo-Croatian scholars as two distinct *standard variants* by the mid-1960s

## 5.2 From lexical to grammatical

The various practices as to how foreign linguistic matter is accepted and adapted into standard Croatian and standard Serbian are strong **differential markers** and inevitably present in any comparative description of Serbian and Croatian. At the same time as foreign influence, especially the influence of English, is viewed as a diversifying factor between Serbian and Croatian, it is also an expression of a globalisation which breaks down borders and homogenises structures around the world. In other words, in the Serbian-Croatian (linguistic) relation the effects of globalisation are symbolic of two contradicting tendencies: heterogenisation vs. homogenisation.

Therefore, I chose to concentrate my empirical research on **contact-induced linguistic innovations**. As I was (and am) not only interested in revealing in which ways contact-induced innovation is implemented in the usage norms of Croatian and Serbian but also wished to discover which general type of innovation is more likely to be accepted by the language users, I chose to focus my research on four types of contact-induced innovations, namely **replication of lexical matter** (in nouns, nominal compounds and semiwords), **word-formational patterns** (in nominal compounds and semiword-compounds), **grammatical matter** (affixes) and **grammatical patterns** (indefinite marking of NPs).

I here present eight conclusions (a-h) concerning contact-induced innovations in Serbian and Croatian and their acceptance in the prescribed norms and usage norms of the two languages:

### **a) Only replicated lexical matter may serve as a differential marker between Serbian and Croatian usage norms.**

When the replicated piece of lexical matter is a noun, my findings based on searches of 61 pairs of **nouns** in the **Croatian National Corpus** (HNK) and the **Corpus of Contemporary Serbian** (SrpKor) make evident that the symbolic value of using a distinctively Croatian word is so high that not only are English and German loanwords less used than their Croatian equivalents but even older and more entrenched **replicas** from French and the Classical languages are being used much less than their **domestic equivalents**. And replicated nouns are in general used less in Croatian than in Serbian.



When the replicated piece of lexical matter was one of the nouns in a **[NUNDECLN]N-compound**, my findings were not quite as unambiguous. The number of different compounds with each of the 18 queried nouns was generally somewhat lower in HNK than in SrpKor. I interpret this difference not as a reluctance in Croatian language users to use the replicated word-formational pattern **[NUNDECLN]N** but rather as their reluctance to use the replicated matter involved. In short, replicated nouns whether used independently as free, declinable nouns or as head or non-head in a **[NUNDECLN]N-compound**, are not as present in the Croatian usage norm as they are in the Serbian usage norm.

So, the efforts by the agents of Croatian language planning through **corpus planning** (normative works and official recommendations), **acquisition planning** (e.g. the mass media and education) and **prestige planning** (admired and prestigious persons and media who choose to use the prescribed norm) which were aimed at enhancing the differences from Serbian in the usage norm of Croatian by demoting foreign lexical matter have been successful, mostly when free lexical items are concerned and less so if the otherwise free lexical item is bound in a compound.

Among the Serbian linguistic authorities the use of domestic words is also recommended over foreign, but only when it is 'possible or sensible to do so' and, there are no objections to lexical matter *per se* as long as it is used 'with moderation and in a sensible way'<sup>293</sup>. So, not only according to the usage norm in the publicistic functional style of Serbian but also according to the prescribed norm, replicated matter is more acceptable in Serbian than in Croatian.

#### **b) Grammatical matter as replicated affixes and replicated bound lexical semiwords do not serve as differential markers in the usage norms of Croatian and Serbian.**

My findings based on queries of 30 **affixes and semiwords** (all but one replicated) in the HNK and SrpKor show that the Croatian and Serbian usage norms do not differ in this respect. The usage norms of the two languages are much more aligned concerning the use of replicated semiwords and the grammatical matter of affixes than they are when the lexical matter of replicated nouns and replicated nominal compounds is

---

<sup>293</sup> For the full quotes from Brborić et al. (2006) and Čupić (1996), cf. p. 162 and 152, resp.

concerned. It also shows us that the Croatian prescribed norm, according to which replicated affixes should be substituted with domestic equivalents, has not had any significant effect on the usage norm. In other words, the Croatian language users do not heed the recommendation from Croatian linguistic authorities to substitute replicated affixes. The similar Croatian and Serbian usages of replicated grammatical matter and dissimilar usages of lexical matter make it evident that it is more difficult to convince language users to substitute bound morphemes and grammatical matter of affixes and semiwords than it is to convince them to substitute lexical matter. Lexical matter, 'content words', seem much easier to manipulate, probably because of their salience as to the linguistic awareness.

**c) The replicated word-formational pattern of [NUNDECLN]<sub>N</sub>-compounds and the grammatical pattern used in indefinite marking of NPs by adding *jedan* or *neki* do not serve as differential markers between the Croatian and Serbian usage norms.**

In my research of [NUNDECLN]<sub>N</sub>-compounds I found that the word-formational pattern had been replicated, but as mentioned above, my findings indicate that this word-formational innovation is more or less reserved for compounds consisting of at least one replicated noun.

Nominal compounding in Serbian and Croatian seem to be limited to **endocentric compounds** and out of the three possible endocentric compound structures outlined in *The Oxford Handbook on Compounding*: coordinate, attributive and subordinate, my findings suggest that only **attributive** and **subordinate** replicated [NUNDECLN]<sub>N</sub>-compounds exist in Serbian and Croatian. Nominal compounds, which in the model language are either exocentric or endocentric coordinate are therefore in Serbian or Croatian transformed into either appositions or endocentric attributive or subordinate compounds.<sup>294</sup>

In contrast to the different attitudes of Serbian and of Croatian linguistic authorities towards replicated matter, these authorities are much more aligned in their objection to the replicated word-formational pattern: [NUNDECLN]<sub>N</sub>. In codification works, such compounds are called 'semi-compounds' (*polusloženice*) and both the linguistic authorities of Croatian and of Serbian recommend their transformation into different kinds of NPs. However, despite the demotion of [NUNDECLN]<sub>N</sub>-compounds

---

<sup>294</sup> Cf. subsection 4.4.1 for discussion and examples

in the prescribed norms, it seems to be a very productive word-formational pattern in both Croatian and Serbian usage norms. Nominal compounds are produced by either combining two individually replicated nouns as in *filter čaj* or *tampon zona* or by combining a replicated noun as the undeclined E1-noun with a domestic noun as E2: in *internet dobavljač*, *kontakt točka* or *kamp-kućica*.

The productivity of this word-formational pattern is presumed to be the result of pattern replication from English. However, similar nominal compound structures have been attested in Serbian and Croatian for centuries and are also believed to be the product of contact, though not with English.<sup>295</sup> Therefore this word-formational pattern is not new to the common **linguistic system** of Croatian and Serbian and should instead be seen as a **use pattern rising from a minor to a major** one in the usage norms of Serbian and Croatian.

My research of the determiners *jedan* and *neki* as possible markers of **indefiniteness** led me to conclude that in the usage norms of Croatian and Serbian they do indeed mark indefiniteness. Simultaneously, however, they mark the presence or lack of specificity of the determined NP. As opposed to the other investigated phenomena, there had been no normative objections to the **article-like use** of *jedan* and *neki* since the beginning of the 20<sup>th</sup> century. Their article-like use is equally present in the Croatian and Serbian usage norm, so my research of this phenomenon did not reveal any difference between the prescribed and usage norms of the two languages. The innovation is no less interesting as it, along with the word-formational innovation described above, confirms my hypothesis that changes in the linguistic system of the two languages are present in both languages. The fact that this has been accepted (or rather: not objected to) by normativists makes it evident that contact-induced linguistic innovations which do not contain any replicated matter are more easily accepted not only into the usage norm, but also into the prescribed norm. The fact that the grammatical category definiteness is not obligatorily marked in Serbian and Croatian shows that a replication of a full pattern of an article-language such as Greek, Turkish, Italian, German<sup>296</sup> or English<sup>297</sup> has not happened. Whether it ever will, is unknown

---

<sup>295</sup> Ancient Greek via Church Slavonic, Turkish, German and Czech, cf. section 4.4 for examples.

<sup>296</sup> Earlier languages of administration or religion in the Serbian and Croatian language area

<sup>297</sup> The dominant global language of today

but some scholars claim to have observed a general **rise of articles**. What is certain, in my view, is that *jedan* and *neki* are very often used to mark indefiniteness and, according to the existing research, they are used more and more in this manner. This increased use makes it yet another possible instance of **a rise from a minor to major use pattern** and this does constitute a change of the shared linguistic system of Serbian and Croatian, so:

**d) The pattern replication of certain features of indefinite articles on to *jedan* and *neki* does change the structural build-up of the languages as does the rise of the word-formational pattern [NUNDECLN]<sub>N</sub>**

### 5.3 Symbolic and communicative value

One of the questions I set out to answer is whether the investigated linguistic innovations bring Croatian and Serbian further from or closer to each other. My *hypothesis* was that the linguistic effects of globalisation would unite rather than divide Croatian and Serbian.

Language's primary *raison d'être* is communicative but at the same time it also has symbolic value. In other words, a language user (speaker) signals something about his identity through his language use. The language is a system with limits. The speakers mostly operate within these limits but when confronted with systems (languages) with other lexical, word-formational and grammatical possibilities, speakers pick up some of these possibilities and apply them in their usage of their own language. If these individual innovations are actualised by many users (or 'cohorts of speakers'<sup>298</sup>), they become part of the **usage norm**.

The usage norm is in this study defined according to the *Prague School* and *Coseriu* who defined the norm as representing what is **acceptable** usage whereas the system represents what is **possible**.<sup>299</sup> Dependent on whether the innovation is one of linguistic matter or pattern, the innovation either is adapted to the system or the system adapts to the innovation. In other words, when pattern is replicated, it integrates with the receiving system thereby altering its limits, but when matter is replicated, it is transformed so that it complies with the receiving system.

---

<sup>298</sup> As Andersen (2008) terms it. Cf. p. 116 for full quote.

<sup>299</sup> cf. subsection 2.2.1 for discussion and citing.

Some users, the **normativists**, wish to steer the development of the usage norm both for communicative and for symbolic reasons. For this purpose an ideal norm for language use is described and prescribed – i.e. the standard language. This study shows very clearly that **replicated linguistic matter has a higher symbolic value than replicated linguistic pattern** for both speakers and normativists. So, when Croatian normativists chose to recommend changes concerning usage of lexical and grammatical matter (for symbolic reasons) they had better chances of changing the norm *because* of the symbolic value. I.e. the proposed changes symbolised values that the majority of users were sympathetic to, namely distancing themselves from everything Serb or Yugoslav. Therefore the Croatian normativists succeeded in ridding not only the prescribed norm but also the usage norm of undesirable lexical matter. So:

**e) Contact-induced replicated lexical matter brings Serbian and Croatian *further away* from each other.**

The grammatical matter of the investigated affixes and the word-formational and grammatical patterns of the nominal compounds and the indefinite marking, however, represent the opposite. Despite the symbolic value of replicated affixes, semiwords as well as the replicated word-formational pattern in the nominal compounds, and despite the recommendations from both Croatian and Serbian linguistic authorities, all these innovations as well as the indefinite marking are equally accepted in the usage norms of Croatian and Serbian. So:

**f) Contact-induced replicated grammatical matter, word-formational patterns and grammatical patterns bring Serbian and Croatian *closer* to each other.**

The reason for this contradicting tendency lies, I believe, in each innovation's **degree of entrenchment** in the linguistic system shared by Croatian and Serbian. The more grammatical and functional an innovation is, the more it is integrated with the linguistic system of the replica language and therefore more deeply entrenched. If the usage of a language has provoked a change of its underlying structure, the linguistic system, and made it normal to form nominal compounds and semiword-compounds in a particular manner or, has made it normal to explicitly mark a NP as indefinite, it is less likely that language planning

efforts will lead to a reversion of this change in anything but the prescribed norm. The communicative value of the possibility to make [N<sub>UNDECL</sub>N]<sub>N</sub>-compounds seems to far outweigh the symbolic value of abiding by the prescribed and recommended rules for combining two nouns. And, the communicative value of using internationally well-known affixes and semiwords is also appreciated by many language users of Croatian even though they lack the “croatianness” of the proposed substitutes. Because of the communicative value of these innovations, speakers of both Serbian and Croatian have accepted them and use them and thus the linguistic effects of globalisation unite Serbian and Croatian.

Another, connected, reason for the resilience of replicated patterns is the general unawareness of its foreign origins. In other words, when the linguistic matter in a replicated pattern is domestic, its symbolic value is not necessarily negative. This means that in a language usage norm which has been successfully adapted to new socio-political circumstances there is much awareness of **overt** contact-induced changes in which a foreign influence (replicated matter) is easily recognisable. On the other hand, a contact-induced change of a pattern is **covert** as it is not necessarily recognised as foreign influence by the average language user, if recognised as an innovation at all. And when it is not recognised as foreign influence there is no positive symbolic value in refraining from using said contact-induced linguistic innovation. So, we may conclude that:

**g) The impact of linguistic authorities on the usage norm is limited to overt changes**

Finally, the answer to the overarching research question would be:

**h) The linguistic effects of globalisation do not create any systemic differences between Croatian and Serbian. Croatian and Serbian react identically to systemic changes induced by language contact. Only on the level of lexical matter do the reactions to the linguistic effects of globalisation enhance the differences between Serbian and Croatian usage norm.**



## Appendices

### Appendix A

In this appendix I have listed all the types of items which I found combined with nouns in  $[[N_{UNDECL}N]_N$ -compounds. In other words, the numbers in Tables 4.4.1, 4.4.7 and 4.4.8 reflect the nouns listed here.

Table 4.4.1 - The nouns queried as element 1 in  $[N_{UNDECL}N]_N$ -compounds.

E1-noun	no. of different E2s		Translation
	SrpKor	HNK	
internet	384	184	<i>Internet</i>
auto	149	134	<i>car</i>
biznis/business	108	23	<i>business</i>
šoping/shopping	30	18	<i>shopping</i>
kontakt	27	25	<i>contact</i>
parking	24	16	<i>parking-lot</i>
kamp	13	10	<i>camping</i>
sladoled	4	1	<i>ice-cream</i>
menadžment/management	1	4	<i>management</i>

Table 4.4.7 – Queried replicated E2-nouns in nominal compounds

E2-noun	no. of different E1-nouns		Translation
	SrpKor	HNK	
zona	22	5	<i>zone</i>
paket	17	6	<i>package</i>
šou/show	17	10	<i>show</i>
biznis/business	9	2	<i>business</i>
koktel/cocktail	1	0	<i>cocktail (-party)</i>
parket	0	2	<i>parquet floor</i>

Table 4.4.8 – Queried domestic E2-nouns in nominal compounds

E2-noun <sup>1</sup>	no. of different E1-nouns		Translation
	SrpKor	HNK	
muzika/glazba	32	27	<i>music</i>
čaj	3 (+6)	1	<i>tea</i>
sladoled	0	0	<i>ice-cream</i>

---

<sup>1</sup> Besides *sladoled* and *glazba* I have decided also to categorise *čaj* and *muzika* as domestic nouns because they are examples of the originally replicated nouns that can be considered to be *adoptees*, i.e. nouns which not even purist normativists think should be substituted with a domestic equivalent. The noun *muzika* is, however, considered foreign in Croatian, which is why I queried both *muzika* and *glazba* in both corpora.



*auto as E1 in a [N<sub>UNDECL</sub>N]N - compound*

**149 types of E2 in SrpKor**

alarm	garaža	mafija	reli
atlas	gas	magazin	revija
bar	gigant	magistrala	sajam
baza	grdosija	mehaničar	salon
beton mikser	groblje	mehanika	saobraćaj
bioskop	grupa	mešalica	sedište
biznis	grupacija	motor	sektor
blokator	guma	obilaznica	servis
boja	hauba	oglasnik	sijalica
bol	helikopter	oprema	simbol
bomba	hit	osiguranje	sirena
bravar	industrija	otpad	škola
brci	instalacija	oznaka	šofer
brend	izlaz	parada	šou
brisač	izložba	park	spektakl
bum	jedinica	parking	sport
buvljak	kabal	patrola	staklo
centar	kamp	perionica	stik
cesta	karoserija	pijaca	stop
četa	karta	pista	takmičenje
cisterna	katalizator	plac	taksi
čistilica/čistalica	klaster	ploča	tradicija
dan	klub	ponuda	transport
delovi	kolona	preporuka	trka/trke
diler	kompanija	pretakalište	trkač
direktor	komplet	prevoz	trkanje
dizalica	komponenta	prevoznik	tržište
dizaličar	komponentaš	prikolica	ubica
dogadjaj	koncern	privreda	udes
električar	kozmetika	proizvođač	ulaz
elektrika	kredit	puk	ulje
elektro opreme	kuća	punilište	uređaj
elektroničar	kultura	put	voz
elektronika	lak	radionica	vozač
ergela	lakirer	radnik	vožnja
esnaf	lift	radnja	
fešta	limar	regija	
firma	limarija	reklama	

### 134 types of E2 in HNK in compounds with *auto* as E1

akustika	košara	patrola	taksi/taxi
alarm	kozmetika	perač	taksist
biznis	kradica	pija/pijac	tehnika
boja	kredit	plin	termin
bomba	kriminal	poduzeće	transformator
branša	kriminalac	poduzetnik	tranzicija
časopis	kuća	poligon	trasa
centar	kućica	pometnja	trka
cesta	lak	populacija	trkalište
cisterna	lakirer	povezivanje	tržište
čistilica	lim	praktikum	tržnica
debi	limar	praonica	turist
destinacija	limarija	presvlaka	tvrtka
dijelovi (pl.)	ljestve (pl. tant)	prijevoz	utrke
diler	lobi	prijevoznik	vlak
div	mafija	prijevoznništvo	vožnja
dizajner	magazin	prikolica	zajam
dizalica	maketa	prilog	
dom	marka	proizvodnja	
električar	mehaničar	put	
elektrika	mehanika	radio	
elektronika	miješalica	radio-kasetofon	
fijaker	mikser	radionica	
gigant	model	rally/reli	
gril	modelar	revija	
guma	modelarstvo	sajam	
hladnjač	muzej	salon	
industrija	naježda	sat	
kamp/camp	nasilnik	servis	
kancelar	navigacija	show	
karijera	navlaka	simulator	
karta	nesreća	sjedalica	
kasko	novinar	škola	
kazetofon	ophodnja	smečar	
kino	oprema	sport	
klub	osiguranje	stop	
kočenje	osiguravatelj	stopiranje	
korpa	otpad	stručnjak	
korporacija	parking	struka	

***biznis/business as E1 in a [N<sub>UNDECL</sub>N]<sub>N</sub> – compound***

**108 types of E2 in SrpKor**

**23 types in HNK**

administracija	klub	pristup	apartman
alijansa	kompanija	proces	centar
analiza	konferencija	program	ciklus
aplikacija	kongres	projekat	elita
areni	konkistador	prostor	establishment
asocijacija	korisnik	putnik	forum
aspekt	korporacija	rečnik	ideal
barijera	kredit	registar	kartica
centar	krug	rešenje	klasa
četvrt	kula	rijaliti šou	liga
dama	lider	rubrika	ljevica
deo	liga	ručak	logika
dijalog	lobi	sala	milje
direkcija	logika	saldo	plan
duh	menadžer	samit	priležnik
džet	menadžerka	šansa	projekt
elita	menadžment	scena	putnik
fakultet	mišljenje	sedište	sektor
fijasko	model	sekcija	središte
forum	mreža	sektor	struka
genij	obračun	škola	svijet
glavar	obuka	specijalitet	utakmica
ideja	odeljenje	starešina	zajednica
imidž	odmor	strategija	
inicijativa	okruženje	struktura	
inkubator	operacija	trend	
istraživanje	paket	trener	
javnost	paradigma	trougao	
jedinica	park	turizam	
kanal	partner	utakmica	
karta	plan	zbivanje	
kartica	planiranje	žena	
kategorija	populacija	zgrada	
klan	poslovica	zona	
klasa	pravilo	žurka	
klima	prijatelj	žurnal	

***biznis/business as E2 in a [N<sub>UNDECLN</sub>]<sub>N</sub> - compound***

**9 types of E1 in SrpKor**

**2 types in HNK**

internet	rentakar	seks	taksi		internet
kazino	ringišpil	šou/show	veb		show
onlajn					

***čaj as E2 in a [N<sub>UNDECLN</sub>]<sub>N</sub> - compound***

**3 types of E1 in SrpKor**

**1 type of E1 in HNK**

kleka
filter
tonik

aroma
-------

*internet as E1 in a [NUNDECLN]<sub>N</sub> - compound*

**384 types of E2 in SrpKor**

adresa	čovjek (ljudi)	glas	katalog
agencija	čvorište	glaslač	kazino
akcija	debata	glasanje	kičma
aktivista	deo	glasilo	kiosk
aktivnost	dimenizija	gledalac	kladenje
anketa	direktorijum	govor	kladionica
aparatus	diskusija	građanin	kladioničar
aplikacija	distribucija	grana	klub
apoteka	distributer	haker	knjižara
arena	dizajn	heroj	kolunikacija
arhiv	dnevnik	hit	komentar
aukcija	doba	hronika	komocija
bajka	dobavljač	igra	kompanija
banka	doktor	igrač	komunikacija
bankarstvo	dokumentacija	igrice	konekcija
baza	dolar	iluzija	konferencija
beduin	domen	ime	konkurs
berza	donacija	impuls	korišćenje
bezbednost	druženje	indeks	korisnik
bibliografija	dućan	informacija	korporacija
biblioteka	džihadista	informisanje	krađa
bilten	džin	infrastruktura	kredit
biznis	ekonomija	intervju	kriminal
blog	ekspanzija	istraga	kriminalac
broj	eksperat	istraživanje	kuća
brouzer	enciklopedija	izdanje	kuloar
budućnost	era	izdavač	kultura
buvljak	erotika	izvor	kupac
čarolija	eskperiment	jugonostalgicar	kupovina
časopis	fan	junak	kurs
centar	fantazija	kabl	kutak
četovanje	film	kafe	kviz
činjenica	firma	kafić	legenda
čitač	forma	kampanja	lingvistika
čitaonica	forum	kanal	linija
čitateljka	galerija	kandidat	link
čošak	generacija	kapacitet	list
čoše	gigant	kartica	ljubavnik

### 384 types of E2 in SrpKor with *internet* as E1

(continued from previous page)

lokacija	paket	prečica	radio
lutrija	parazit	predator	rang-lista
magazin	park	predstavljanje	raskolništvo
magistrala	parlament	preduzeće	rasprava
manifest	paškvilo	preduzetnik	rat
manijak	pedofil	pregled	razgledanje
marketing	pedofilija	preljuba	rečnik
mediji	performans	premijer	redakcija
menadžer	pesma	prepiska	reference
milioner	peticija	pretplata	referendum
modem	pevač	pretraga	registar
mornar	pijaca	pretraživač	reklama
mreža	piraterija	prevara	rekorder
nadzornik	pisac	prevarant	represija
nalog	pismo	prevodilac	rešenje
navigacija	plaćanje	prezentacija	resurs
nedeljnik	platforma	priča	revolucija
novinar	podatak	pričaonica	rudnik
novinarstvo	podrška	prijatelji	ruter
novine	pojam	prijava	sadržaj
oblak	poker	prijenos	sagovornik
odeljenje	policajac	priključak	sajt
odrednica	policija	pristup	sala
oglas	politika	prisustvo	šaljivdžija
oglašavanje	ponuda	prodaja	saobraćaj
oglasnik	populacija	prodavnica	saobraćanje
okean	poredak	profesionalac	sastajalište
okruženje	pornografija	program	sat
okupljanje	portal	projekat	saučešće
onlajn štednja	poruka	prosjak	savet
operator/operatorer	posao	prostor	savetnica
opslužilac	poseta	protokol	savetovalište
opština	posetilac	proza	segment
organizacija	poslovanje	prozor	sekretarica
orkestar	posrednik	publika	senzacija
oružje	pošta	publikacija	serijal
ovisnik	potvrda	računar	server
oznaka	pravac	rad	servis

### 384 types of E2 in SrpKor with *internet* as E1

(continued from previous page)

sesija	telefon	zemljotres
signal	telefonija	zlostavljanje
simfonija	telefoniranje	zavisnik
singl	televizija	zavisnost
sistem	teoretičar	zloupotreba
skupština	terminal	znanje
slava	terminologija	znatiželjnik
sloboda	trag	zvezda
softver	trener	
šop	trgovac	
šoping	trgovina	
špijunaža	tribina	
sposobnost	trka	
sredstvo	tržište	
standard	turnir	
stanica	učenik	
statistika	učionica	
štednja	udruženje	
stil	ugovor	
strana	umetnost	
stranica	umreženost	
striptizeta	univerzum	
stručnjak	upoznavanje	
struktura	usluga	
studija	veb sajt	
stvari	velesila	
stvarnost	verzija	
supermarket	vest	
surfer	veza	
surogat	vic	
suvenirnica	video	
svakodnevnica	vreme	
tabla	zabava	
tajkun	zajednica	
takmičenje	zakonitost	
tehnologija	žargon	

### 184 types of E2 in HNK with *internet* as E1

adresa	forum	manijak	poslovanje	servis
afeta	galerija	mediji	poslužitelj	site
agencija	generacija	mjesto	potencijal	šoping
akcija	gigant	moćnik	povezanost	stanica
analitičar	glasanje	mreža	praksa	štrajk
anketa	glasilo	nalog	pravo	stranica
apel	glasovanje	natjecanje	preglednik	strategija
aplikacija	informacija	netječaj	pretraživač	studij
arhiv	infrastruktura	news-grupa	prezentacija	sustav
aukcija	instalacija	neznalica	pričaonica	svakodnevnica
balet	izdanje	novine	prijenos	tehnologija
banka	kafić	odjel	priključak	telefonija
bankarstvo	kategorija	odnos	prioritet	tražilica
baza	kiosk	odsjek	pristup	trend
biznis	klađenje	ofenziva	prodaja	trgovac
bonus	kladionica	ovisnik	prodor	trgovina
boom/bum	klub	pećnica	produkcija	tržište
broker	knjižara	pedofilija	proizvod	tvrtka
cafe, caffe, kafe	književnost	performance	projekt	ugovor
čajana	kockanje	peticija	promet	umjetnost
časkanje	kockarnica	pilot-projekt	promidžba	umreženje
časopis	kompanija	pismo	promocija	usluga
centar	komplet	pitanje	promotori	uvjet
cijena	komuniciranje	plaćanje	prostor	veza
čovjek (ljudi)	komunikacija	podatak	protokol	vrijeme
dionica	konekcija	podružnik	prozor	web stranica
disident	konferencija	poduzetnik	publika	web sučelje
div	konzultacija	polemičar	publikacija	zajednica
dnevnik	konzultant	policajac	rad	zaljubljenik
dogadjaj	korisnik	policija	radio postaja	zarada
domen	kriminalac	polufinalist	radio stanica	zbornik
dostignuće	kutak	ponuda	radionica	znalac
dućan	linija	populacija	redakcija	
dvomjesečnik	list	pornografija	roaming	
edukacija	ljubavnik	portal	sadržaj	
ekonomija	lobi	poruka	sajam	
e-mail	magazin	posjet	sajt	
festival	manifestacija	posjetitelj	sektor	



***kamp as E1 in a [N<sub>UNDECLN</sub>]<sub>N</sub> – compound***

**13 types of E2 in SrpKor  
in HNK**

izlet	naselje
izletište	oprema
jedinica	prikolica
kamion	smeštaj
kuća	turista
kućica	žurka
lider	

**10 types of E2**

jedinica	oprema
kuća	parcela
kućica	prikolica
mjesto	prostor
naselje	vozilo

***koktel/cocktail as E2 in a [N<sub>UNDECLN</sub>]<sub>N</sub> – compound***

**1 type of E1 in SrpKor**

**no E1 in HNK**

rum
-----

***kontakt as***

***E1 in a [N<sub>UNDECLN</sub>]<sub>N</sub> – compound***

**27 types of E2 in SrpKor  
in HNK**

adresa	linija
ambasada	mejl
biro	oglas(i)
brava	osoba
centar	podaci
emisija	program
hipoteza	šou-program
igra	sport
improvizacija	staratelj
informacija	tačka
kabl	telefon
kartica	zemlja
ključ	zona
krug	

**25 types of E2**

adresa	osoba
brava	podaci
bravica	policajac
broj	policajka
burza	policija
centar	program
emisija	rajon
igra	sastanak
informacije	telefon
ključ	trening
kopija	veza
lista	zemlja
mogućnost	

***menadžment/management as E1 in a [N<sub>UNDECLN</sub>]<sub>N</sub> – compound***

**1 type of E2 in SrpKor 4 types of E2 in HNK**

izum
------

direktor	program
plan	sustav

***muzika/glazba as E2 in a [NUNDECLN]<sub>N</sub>- compound***

**32 types of E1 in SrpKor**

**27 types of E1 in**

**HNK**

big bend	latino
blugras	lingala
bluz	motaun
dens/dance	pačanga
džez	pank
fado	pleh
fank	pop
folk	rege
gospel	rep/rap
haus	rok
hevi-metal <sup>2</sup>	rokabili
hiphop	rokenrol
indi	salsa
kantri	soul
kičera <sup>2</sup>	šund
klub	trans

a capella	metal
afro-pop	pleh
blues	pop
bombon	pop-rock
country	punk
dance	rap
drum'n'bass	rock
flamenco	rock'n'roll
folk	surf-rock
funk	trip hop
gospel	western
heavi metal	
hip-hop	
hit	
jazz	
latino	

***paket as E2 in a [NUNDECLN]<sub>N</sub>- compound***

**17 types of E1 in SrpKor**

**6 types of E1 in**

**HNK**

ADSL	ekonomija	standard		azitromicin
ambijent	IMS	start		DVD
biznis	Internet	trend		miks
BTS	Kosovo	velnes		poklon
DNK	poklon	vikend		softver
DVD	prednost			start

***parket as E2 in a [NUNDECLN]<sub>N</sub>- compound***

**No E1 in SrpKor**

**1 type of E1 in HNK**

lamel
-------

<sup>2</sup> Also spelled *hevi metal* and *hevimetal*

***parking as E1 in a [N<sub>UNDECLN</sub>]<sub>N</sub> – compound***

**24 types of E2 in SrpKor**

časovnik	plac
čvor	površina
garaža	pozicija
kapacitet	prostor
karta	sat(ovi)
kartica	senzor
kazna	servis
kolaps	sistem
kontrola	služba
kontrolor	tarifa
mesto	usluga
metar	zona

**16 types of E2 in HNK**

automat	tvrtka
broj	udruga
karta	usluga
kartica	zona
koncesionar	
mjesto	
oprema	
položaj	
površina	
prostor	
sat(ovi)	
senzor	

***sladoled as E1 in a [N<sub>UNDECLN</sub>]<sub>N</sub> – compound***

**4 types of E2 in SrpKor**

bar	kolač
duga	torta

**1 type of E2 in HNK**

pakiranje
-----------

***šoping/shopping as E1 in a [N<sub>UNDECLN</sub>]<sub>N</sub> – compound***

**30 types of E2 in SrpKor**

agencija	navika
amater	pohod
asistenat	program
centar	put
destinacija	servis
festival	sezona
globalizacija	spot
grad	torba
groznica	tura
hotel	turista
hram	ulica
izlog	vikend
kultura	voz
lista	zabava
maraton	zona

**18 types of E2 in HNK**

centar	turist
groznica	turizam
izlet	usluga
izletnik	
lanac	
lista	
ludilo	
meka	
ponuda	
putovanje	
špica	
središte	
stanka	
stranica	
tura	

***šou/show as E2 in a [N<sub>UNDECLN</sub>]<sub>N</sub>- compound***

**17 types of E1 in SrpKor**

auto	pop
blokbuster	radio
gala	reli
hiphop	rok
hit	slajd
kontakt	talent/talenat
košmar	tango
kviz	veb
multitalent	

**10 types of E1 in HNK**

auto	tv bingo
cocktail	
Internet	
jazz	
kviz	
laser	
parlament	
talent	
tv	

***zona as E2 in a [N<sub>UNDECLN</sub>]<sub>N</sub>- compound***

**22 types of E1 in SrpKor**

bikini	minus	talenat	
biznis	monitoring	tampon	
Dunav	parking	tunel	
evro	Šengen	UNPA/Unpa	
katran	SFN	vikend	
konkakaf	šoping	VIP	
kontakt	tabu	zombi	
miks			

**5 types of E1 in HNK**

euro
miks/mix
parking
tampon
vikend

## Appendix B

In this appendix I have listed all the types of items which I found combined with the selected preposed semiwords and prefixes in nouns that met the criteria for my searches. In other words, the numbers in Table 4.5 below reflect the number of nouns (in E2-position) found in combination with the queried preposed elements (E1) listed on the following pages.

Table 4.5 – Queried preposed semiwords and prefixes

Preposed semiwords	no. of different E2s		
	SrpKor	HNK	
art	66	57	
akva/aqua	7	5	
jugo	68	141	
aero	26	18	
avio	114	53	
eko	121	204	
etno	133	89	
bruto	62	68	
Prefixes	SrpKor	HNK	Recommended substi- tutes in HJS <sub>302</sub>
hiper	57	31	-
eks/ex	30	53	<i>bivši</i> <sub>303</sub> , <i>raz</i>
kontra	142	87	<i>protu</i>
kvazi	156	129	<i>laži, nadri, nazovi</i>
para	36	34	<i>laži, nadri, nazovi, ne, pa, polu</i>
re	134	141	<i>ponovni</i>
ko <sub>304</sub>	53	25	<i>su-</i>

---

<sup>302</sup> In The Croatian Language Manual (Hrvatski jezični savjetnik) Barić et al (1999) recommend a substitutes for the replicated prefixes, as listed here. *hiper-* is not included as a prefix in the HJS.

<sup>303</sup> Except the adjectives *bivši* ('former') and *ponovni* ('repeated'), all the recommended substitutions are prefixes

<sup>304</sup> There were no nouns, meeting the criteria, in either corpus, with the prefix *ko-* spelled *co-*.

**aero as E1****26 Nouns in SrpKor**

defile	snimanje
foto snimak	temperatura
fotografija	transport
kapacitet	tunel
klub	zagađenje
kočenje	zagađenost
kompanija	zagađivač
koncern	zamagljivanje
linija	
miting	
mobil	
model	
modelarstvo	
perspektiva	
prostor	
put	
reli	
sistem	

**18 nouns in HNK**

baza
fotograf
fotografija
gluhoća
kickboxing
klub
miting
model
modelar
modelarstvo
poezija
polucija
slikarstvo
snimak/snimka
tehnika
transport
triangulacija
tvrtka

**akva or aqua as E1****7 nouns in SrpKor**

aerobik
park
ski
skijaš
skuter
teatar
terapija

**5 nouns in HNK**

aerobic/k
centar
park
promet
sport

**art as E1****66 nouns in SrpKor**

aktivizam	lokal
album	mašina
atelje	menadžer
bioskop	meta
broker	objekat
centar	<b>o</b> grafija
delo	<b>o</b> teka
diler	paviljon
direktor	performans
dvorana	pop
fan	portret
fest	prevara
festival	produkcija
film	profesionalac
finale	program
fondacija	projekat
forma	prostor
fotografija	radionica
galerija	rok
grupa	šara
hepening	scena
hiper-elita	sektor
ikona	simptom
instalacija	sistem
intervencija	slikar
izgled	slikarstvo
izložba	teoretičar
kanal	terapeut
kino	triler
kolaž	umetnik
komad	vandalizam
konsultant	vid
lavrint	zona

**57 nouns in HNK**

caffe	objekt
centar	<b>o</b> ljubaca
cirkuserija	ostvarenje
dealer	<b>o</b> teka
design	poetika
direktor/director	ponuda
diskont	predstava
domet	produkcija
ekscs	program
festival	projekt
fikcija	prošlost
film	prostor
filmaš	provokacija
forma	publika
fotografija	rock
funk	scena
galerija	senzor
glazba	slagalica
hit	strip
kafić	turizam
kino	umiranje
klub	video
kolega	vikend
kompromis	vodič
koncept	
krv	
list	
ljubić	
magazin	
manir	
menadžer	
menedžment	
net	

The linking element -o- between *art* and the noun in E2-position is written in bold.

**avio as E1****114 nouns in SrpKor**

agencija	kombinacija	put	udes
akrobacija	kompanija	putnik	usluga
akrobata	komponent	putovanje	varijanta
akrobatika	konstruktor	raketa	veza
aranžman	konzorcijum	rat	vlasti
baza	korporacija	reli	zaštita
benzin	let	sajam	
bezbednost	limar	saobraćaj	
biznis	linija	sedište	
bomba	luka	sektor	
bombardovanje	maketa	servis	
brigada	mehaničar	sfera	
budžet	meni	sistem	
čarter	metoda	škola	
cisterna	model	služba	
dejstvo	modelar	smer	
destinacija	modelarstvo	snimak	
dispečer	mogul	snimanje	
enterijer	most	sport	
ešalon	motor	strada	
firma	mreža	sukob	
flota	navigacija	svetlo	
fobija	nesreća	svrha	
gorivo	obrok	taksa	
grupacija	operacija	taksi	
gubitaš	park	taksiranje	
gužva	partner	tehničar	
industrija	pista	tehnika	
instrument	pošiljka	tehnologija	
inženjer	prevoz	tradicija	
inženjering	prevozilac	transfer	
izlet	prevoznik	transport	
izložba	pristanište	transporter	
karta	projektil	tranzit	
katastrofa	promet	tržište	
klub	puk	tura	



**53 nouns with *avio* as E1 in HNK**

aranžman	kompanija	program
baza	koncern	promet
benzin	konstrukcija	putnik
bomba	let	putovanje
čarter	limar	robot
cisterna	linija	sajam
cluster	maketa	šalter
destinacija	modelar	servis
dolazak	most	sjedalo
elektrika	motrenje	sklop
generator	operater	snimatelj
granata	oprema	stajanka
horizont	paleta	taksi
industrija	park	tehnologija
kapacitet	partner	tretiranje
karta	posjet	tretman
karton	prijevoz	tvrtka
klub	prijevoznik	

**bruto as E1****62 nouns in SrpKor**

bilans	potrošnja
budžet	površina
cena	premija
dinar	prihod
dividenda	primanje
dobit	prinadležnost
dobitak	princip
dohodak	produkt
doprinos	profit
društvo	proizvod
dug	proizvodnja
gubitak	promet
investicija	prosek
isplata	rast
izmirenje	režim
iznos	rezultat
kamat	satnica
kapital	sistem
kilogram	stopa
kolač	teret
količina	težina
marža	tona
masa	tonaža
metar	trošak
naknada	varijanta
naplata	vrednost
obračunavanje	zapremnina
oporezivanje	zarada
osnovica	
otplata	
otpremnina	
plaćanje	
plasman	
plata/plaća	

**68 nouns in HNK**

bilanca	plasman
božićnica	porez
cijena	potraživanje
dimenzija	povećanje
dobit	površina
dobitak	premija
dohodak	pričuv(a)
energija	prihod
gradijent	priljev
gubitak	primanje
honorar	prinos
imovina	profit
investicija	prohod
isplata	proizvod
izdatak	proizvodnja
iznos	promet
kg	prosjeak
kredit	rad
kuna	regresa
kvadrata	rezerva
marža	satnica
masa	smanjenje
minimalac	snaga
mirovina	sredstva
načelo	svota
nadoknada	težina
naknada	tona
nosivost	tonaža
obveza	ulaganje
osnov(a)	uplata
osnovica	utržak
ostatak	vrijednost
output	zaduženje
plaća	zarada

**eko as E1****121 nouns in SrpKor**

agronomija	kamper	oznaka	terijanac
akcija	karta	pakovanje	turizam
aktivista	katastrofa	pare	učionica
ambalaža	kazino	park	umetnost
auspuh	kodeks	patrola	upravljanje
automobil	kompleks	petrol	utopija
autoreli	kontrola	pijaca	venčanje
baza	koža	planiranje	voda
bazar	kozmetika	pokret	vozilo
bilbord	krzno	policajac	vrtić
biznismen	kuća	poljoprivreda	začin
bloger	kultura	porez	zombi
bomba	levičar	posledica	zona
brana	mafija	premaz	
brižnik	marksista	problem	
bunda	materijal	program	
centar	milijaderka	proizvođač	
česma	mit	projekat	
dažbina	mladoženja	propast	
dijamant	mobil	protekcionizam	
dinar	moda	radionica	
dizajn	model	raj	
dizel	modernizacija	razvoj	
društvo	motor	rečnik	
džins	musketar	revolucija	
etika	muzej	rizik	
film	nagrada	robot	
fond	naknada	selo	
fotografija	naziv	skepticizam	
gastronom	nevolja	skeptik	
gorivo	novinar	sredina	
hidrocentrala	oaza	standard	
hrana	obeležavanje	stanište	
industrija	olimpijada	struja	
inovacija	orijentacija	tačka	
kamp	osetljivost	taksa	

## 204 nouns with *eko* as E1 in HNK

akcident	igračka	med	politika	sekcija	tvornica
akcija	imanje	misija	polje	seljak	udruga
aktivizam	incident	mlin	poljoprivreda	selo	udruženje
aktivnost	inspekcija	mogućnost	poljoprivrednik	seminar	ulaganje
anarhista	institucija	motor	ponuda	sjenica	umjetnica
banda	izvor	mreža	porez	škola	urod
benzin	jabuka	muzej	poruka	skrbnik	uvjet
blokada	Jadran	naknada	posada	skup	varijanta
bod	jaje	naselje	poticaj	slagalica	vijest
bomba	kamp	naziv	povrće	slika	vino
brod	kanta	nedoumica	prečišćavanje	slom	vlak
brodić	karta	nesporazum	predmet	smjernica	voće
čaj	katastrofa	nesreća	predstava	soja	vozilo
centar	kiparstvo	njoki	priča	štant	vožnja
certifikat	klub	novinar	prijetnja	standard	vrtić
cisterna	kompleks	obitelj	priznanje	staza	zadruga
deponija	konferencija	obveza	problem	stil	zahod
dio	kontejner	odbor	program	stožer	zamisao
dizel	krava	opasnost	proizvod	strategija	zastava
dogadaj	kriminal	osnova	proizvođač	struja	želja
dovoljnost	kriterij	otok	proizvodnja	studij	znak
efikasnost	kruh	označavanje	projekt	summit	zob
emisija	kukuruz	pakiranje	propis	susret	zona
entuzijast	kulen	palenta	prostor	suvenir	
farma	kultura	pano	prosvjed	svijet	
farmer	kurir	papir	pšenica	taksa	
feminizam	kviz	park	puretina	tema	
film <sup>305</sup>	laboratorij	parola	radionica	teorija	
filter	lanac	patrola	raj	terminator	
fond	linija	piletina	razvoj	terorizam	
fundacija	ljestvica	pilić	renta	test	
gastronom	lobi	plac	repa	testiranje	
gospodarstvo	logotip	plakat	sajam	tim	
grupa	mafija	plastika	salama	tjestenina	
hrana	markica	pojas	sastav	turist	
igra	materijal	pokret	savez	turizam	

---

<sup>305</sup> In the meaning 'clingfilm' or 'plastic wrap' not in the meaning 'moving pictures'

**etno as E1**

**133 nouns in SrpKor**

album	izložba	objek(a)t	simfonija
ambijent	kafe	obrazac	skupina
ansambl	kamp	opera	slučaj
antropologija	kapa	pank	soba
apartman	kapital	park	šop
aranžman	karakteristika	pesma	specijalitet
atomist(a)	karate	pevačica	stanje
bajka	karta	pevanje	stil
baština	klaustrofobija	pijaca	stranka
bend	klupica	ponuda	stvaralaštvo
botanika	komad	pop	svadba
butik	kombinacija	pornografija	teatar
celina	kompleks	posluženje	tradicija
centar	konačište	postavka	trend
demokratija	koncert	pozadina	trka
dešavanje	konsultant	pravac	turizam
detalj	kuća	predmet	varijanta
dogadjaj	kuhinja	predstava	vodič
domaćinstvo	kultura	prezentacija	vrednost
dvorište	kutak	priredba	začin
džez	lokalizam	pristup	zapis
eksponat	manifestacija	pritisak	zbirka
elemen(a)t	menadžer	program	život
fazon	mlekar	proizvod	zvezda
festival	model	radionica	zvuk
film	motiv	restoran	
folk	muzej	roba	
galerija	muzičar	rok	
grupa	muzičarka	sabor	
haljina	muzika	sadržaj	
hip	nacionalist(a)	sajam	
hrana	nacionalizam	salaš	
igra	naselje	sastav	
instrument	nasleđe	selo	
inženjering	nota	serija	
istorija	oaza	simbolizam	

### 89 nouns with *etno* as E1 in HNK

album	kombinacija	simbol
ambijent	koncert	skupina
ansambl	korijen	sound
antropologija	kulisa	spot
atrakcija	lice	stilizacija
band/bend	manir	susret
baština	maska	svadba
detalj	matrica	tema
dizajn	mentalitet	turizam
djelo	milje	usmjerenje
dokumentacija	miris	val
element	mišung	varijacija
festival	motiv	večer
film	muzej	zbirka
folk	nedjelja	zona
genocid	objekt	zvijezda
glazba	park	zvuk
glazbenica	patina	
glazbenik	pjevačica	
grupa	područje	
hit	pokret	
Hrvatska	pop	
ideologija	pratnja	
improvizacija	pristup	
infekcija	prizvuk	
infuzija	pub	
izričaj	radionica	
izvođač	reportaža	
izvođačica	restoran	
izvorište	revija	
jazz	rock	
jazzerica	sadržaj	
karakter	sazvučje	
karavana	scena	
klub	selo	
kolekcija	serija	

**eks or ex as E1****30 nouns E1in SrpKor**

bend	špijun
brat	teritorija
član	Ukrajinc
članica	vladika
frontmen	
gerilac	
gradonačelnik	
guverner	
JU	
Juga/Jugo	
Jugosloven	
Jugoslavija	
kralj	
ministar	
ministarka	
ministrica	
načelnik	
potpredsednik	
predsednik	
premijer	
princ	
prvak	
reprezentativac	
republika	
šampion	
selektor	

**52 nouns in HNK**

priručnik	ljevičar
agent	magistar
bilder	major
bolnica	ministar
bubnjar	miss
car	načelnik
čelnik	nogometaš
banka	novinar
desničar	općina
diktator	pjevač
drug	predsjednik
država	prometnjak
frontman	prvoligaš
galerija	republika
glasnogovornica	restaurant
glazbenik	Riječanin
gradonačelnik	Savez (komunista)
hotel	šef
Juga/Jugo	Šeks
Jugoslavija	teritorij
kancelar	tvornica
kombinat	umjetnik
komunist	uposlenik
komunizam	vlasnik
koncesionar	vlast
košarkaš	YU

**hiper as E1****57 nouns in SrpKor**

aktivnost	rasejanje
ambicija	ravan
bezbednost	reakcija
centralizacija	realista
ciklus	realizam
drama	realnost
država	refleksija
elita	relativizam
emotivnost	saznanje
evropa	seksualizacija
funkcija	seksualnost
generalizacija	senzibilnost
globalista	senzitivnost
individualizam	senzualnost
ironičnost	sila
katolicizam	skup
kič	stvarnost
književnost	svemir
kocka	tehnologija
kockica	veza
koincidencija	život
kontrola	
kriticizam	
kritičnost	
market	
modernista	
modernizam	
moral	
motivisanost	
mozak	
organizacija	
osetljivost	
parametar	
pijaca	
produktivnost	
prostor	

**31 nouns in HNK**

aktivac
aktivnost
ambicioznost
antagonizam
energija
funkcija
historicizam
individualizam
kocka
market
nacionalist
nacionalizam
normiranost
osjetljivost
poetika
političnost
politiziranost
potrošnja
produkcija
produktivac
realist
realizam
realnost
seksualnost
senzibilnost
senzitivnost
sila
stimulacija
stvarnost
trgovina
uspješnost



*jugo as E1*

**68 nouns in SrpKor**

armija	neuroza
asocijacija	nostalgičar
dinar	nostalgičarka
diplomatija	nostalgija
drum	patriotizam
društvo	pesnik
država	podzemlje
ekonomika	pomladak
elemenat	povratnik
faleristika	predsednik
film	projekat
fudbal	prosek
govor	prostor
inspekt	provokacija
jedinstvo	region
jelovnik	republika
komunista	saradnja
komunizam	sfera
koncert	škola
koral	staljinista
kriza	štampa
list	sused
lokalitet	tenk
loto	torba
mafija	trauma
makartizam	trobojka
marksizam	veteran
maršal	vino
mediji	viza
metak	vojska
mornarica	vreme
muzičar	zapad
nacionalista	zbivanje
narod	zvezda

**141 nouns with *jugo* as E1 in HNK**

admiral	izbjeglica	nostalgija	stranka
afera	izvoz	obveza	student
agent	jednoulje	oficir	susjed
agresija	karijerist	ogranak	sustanar
ambasada	kinematografija	okupacija	švaba
aparatčika	klerikalizam	okvir	svijest
apokalipsa	klub	oportunizam	tamnica
armada	komunist	pakiranje	televizija
armija	komunizam	paprika	tragedija
asosijacija	komunjara	parlament	trobojnica
banka	komunjarstvo	partizan	tržište
baza	kontekst	pisac	turizam
birokracija	konzervativac	poduzetnik	tvorevina
boljševik	konzul	policija	tvrtka
boljševizam	konzulat	politika	uniforma
bon	kriza	povijest	unija
brzina	kup	predsjednik	unitarist
carinik	legenda	predstavnik	unitarizam
dinar	liga	premijer	vampir
diplomacija	mafija	prostor	varijanta
dopisnik	mafijaš	punk	viza
država	manjina	ratovanje	vlada
državljanin	marksizam	regija	vlast
ekonomija	masonstvo	registracija	vlastodržac
euforija	miješalica	republika	vojnici
federacija	mornarica	režim	vojska
fil	mrak	režiser	vrijeme
filstvo	nacionalist	rovarenje	zajednica
flota	nacionalizam	savez	zastava
garnizon	nasljedstvo	saveznik	zatvor
general	navijač	shema	zrakoplov
golman	nogomet	sistem	zrakoplovstvo
gost	nogometaš	služba	zvijezda
inercija	nostalgicar	socijalizam	
integralist	nostalgicarstvo	soldateska	
iskustvo	nostalgičnost	stav	

**ko as E1****55 nouns in SrpKor**

autor	načelnik
autorka	operacija
autorstvo	organizacija
direktor	organizator
direktorka	osnivač
distribucija	pilot
domaćinstvo	potpis
egzistencija	predlagač
egzistiranje	predsednica
evolucija	predsednik
finansijer	predsedništvo
finansiranje	producent
gradonačelnik	produkcija
habitacija	produkt
habitant	pronalazač
habitovanje	reditelj
inicijator	rediteljka
izvestilac	režija
kapiten	režiser
kurator	scenarista
kuratorka	scenaristkinja
kustos	sekretar
lider	selektor
menadžer	selektorka
ministar	sponzor
ministarka	transformacija
ministarstvo	

**25 nouns in HNK**

aranžer
autor
autorica
autorstvo
branitelj
edicija
egzistencija
egzistiranje
financiranje
investitor
investitorstvu
ministar
predsjedatelj
predsjednik
producenti
produkcija
realizator
redatelj
redateljica
referat
režiser
scenarist
selektor
sponzor
vladar

**kontra as E1****142 nouns in SrpKor**

adut	krivina	politika	sila
agent	kultura	ponuda	skup
agitacija	kupovina	poruka	slika
akcija	kvalifikacija	potez	smer
analiza	mantra	potraživanje	snaga
apel	marš	pozicija	špijun
argumenat	melodija	praksa	špijunaža
atak	mera	pravac	stav
balans	miting	pravopis	strana
čin	mitingaš	predlog	struja
demonstracija	model	pretnja	stvaranje
demonstrant	nabavka	priča	svet
diplomacija	nacrt	primer	svetlo
dokaz	nadimak	prirubnica	terorizam
dokumentacija	nagib	proces	teza
edukacija	napad	produktivnost	transfer
efekat	nasilje	program	tudica
ekspertiza	net	propaganda	tužba
elita	obaveštajac	protest	tvrdnja
faza	obaveštajka	puč	udar
garancija	objekat	razlog	udarac
garant	odgovor	reakcija	usluga
glas	ofanziva	referendum	ustupak
hvalisanja	okupljanje	reforma	varijanta
igra	opis	reformacija	verzija
imidž	optužba	regulacija	vetar
indikacija	osvajanje	revolucija	vindikacija
inicijativa	osvećenje	revolucionar	vlast
investicija	osvetljenje	rezolucija	voz
isporuka	partija	rezultat	zahtev
izveštaj	peticija	ritam	žalba
kampanja	pisanje	sadržaj	zavera
kletva	pitanje	samit	zaverenik
knjiga	plan	saopštenje	žudnja
koncert	pobuna	scena	
kritika	poduhvat	sekularizacija	

**86 nouns with *kontra* as E1 in HNK**

agresija	osveta	transfer
akcija	pad	tužba
argument	peticija	učinak
argumentacija	pitanje	udar
avangardist	plakat	udarac
blef	plan	uloga
božac	potez	unija
efekt	pozicija	uvoz
favorit	poziv	uzorak
gard	pravac	vatra
igra	prekršaj	vijest
indikacija	prijedlog	vjetar
institucija	prilog	volej
isporuka	primjer	želja
isprika	pritisak	
izvoz	prodor	
jamstvo	produktivnost	
kompensacija	propaganda	
korak	proslava	
kultura	prosvjed	
legenda	reakcija	
lekcija	referat	
lijek	reformacija	
lobiranje	religija	
miting	revolucija	
mjera	revolucionar	
nagib	ritam	
napad	rješenje	
nasilje	sastav	
obavještajac	skup	
obavještajka	smjera	
ofenziva	špijunaža	
oganj	štrajk	
operacija	summit	
oplata	svjedočenje	
optužba	teza	

**kvazi as E1****156 nouns in SrpKor**

agencija	ekspert	kviz	politika	vidovnjak
akademik	elita	legalizacija	pomirenje	vikendaš
akcija	elitizam	legitimacija	posrednik	vinjak
anketa	euforija	lekar	pravilo	vlada
argumenat	fašizam	levica	pravoslavlje	vlasništvo
argumentacija	federacija	levičar	privređivanje	vođa
astrolog	filolog	ličnost	producent	vrednost
atom	filozof	literatura	proizvodnja	zadrugarstvo
autonomija	filozofija	majstor	projekat	zajam
banja	funkcija	manuskript	prostor	zaštita
banka	govor	master	protektorat	znanje
biologizam	guru	matrijarhat	rasprava	zvuk
borac	harizma	memoari	recesija	
bratstvo	heroj	modernitet	reforma	
budžet	hipi	moral	reformator	
citāt	igra	mudrost	revolucija	
članica	institucija	nacionalista	revolucionar	
čuvar	intelektualac	naivnost	sigurnost	
Dedinje	inteligencija	naracija	sniženje	
demokrata	investitor	naučnik	sociologizam	
demokratija	isključivost	naučnost	straživanje	
deonica	ispovest	nauka	stručnjak	
dijagnostikovanje	istina	navijač	student	
dijagram	istoričar	nezavisnost	studija	
dijalog	istoricizam	norma	sud	
disident	istoriografija	ombudsman	suverenost	
diskoteka	izbor	opozicija	tajnost	
dogma	izgovor	patrijarh	teorija	
dokumenat	jednoglasnost	patriota	testament	
domoljub	junak	patriotizam	teza	
doživljaj	komedija	pevač	totalitet	
država	komora	pevačica	trener	
državljanstvo	komuna	pitanje	udžbenik	
državnost	kontrakt	plemić	umetnik	
duhovnost	kristal	policija	uspeh	
ekologija	kultura	političar	veroučitelj	

**129 nouns with *kvazi* as E1 in HNK**

Amerika	hibrid	oporbenjak	stranka
analiza	historičar	ozbiljnost	straživanje
anketa	hrvatstvo	pansion	stručnjak
argument	idealizam	parlament	subverzija
argumentacija	ideologija	patetika	sukob
autoritet	inkarnacija	patriotizam	tajnovitost
biznismen	intelektualac	pedagogija	teorija
borba	intelektualizam	pobjednik	transakcija
božanstvo	internacionalac	poduzeće	tranzicija
braća	kapitalist	poduzetnik	trgovac
car	koalicija	pokajanje	učenost
cinizam	koncert	pokret	udruga
Crkva	konspiracija	političar	umjetnost
demokracija	kršćanin	politikanstvo	ustav
demokrat	liberal	politiziranje	vlasnik
demokratskost	literat	posttranzicija	zabava
desnica	logika	povjesničar	zaplet
desničar	maćeha	poznaničar	zaštitnik
devalvacija	mačizam	prednost	zavodnik
dijalog	materijal	pregovor	znanost
dioničarstvo	menadžer	prijatelj	znanstvenost
diplomacija	mjera	prikaz	
diplomat	moć	primjena	
dokument	model	proizvodnja	
dokumentarac	monarhizam	rasprava	
dokumentarizam	monolog	rat	
domoljublje	monopol	religija	
država	moralist	republika	
duplerica	nacionalizam	revolucija	
ekspert	navijač	rješenje	
elita	neprijatelj	seljak	
europejac	neprijateljstvo	sindikalizam	
federacija	novac	sindikat	
firma	novinarka	sloboda	
formalnost	obitelj	spektakl	
frajer	oporba	Stradun	

*para as E1*

**36 nouns in SrpKor**

carina	kuvarica
diplomacija	legitimitet
država	literatura
državica	organizacija
državnost	patriotizam
esejistika	ratnik
fejltonistika	rukovodstvo
fenomen	struktura
fond	stvarnost
formacija	teatar
grupa	tehnologija
ideologija	tema
institucija	univerzijada
istorija	univerzitet
istoriografija	vlada
književnost	vlasnik
kultura	vojniki
kuvar	vojska

**34 nouns in HNK**

djelovanje	profesionalac
država	profesionalnost
državica	skupština
elita	služba
gerilci	stranka
historija	strategija
institucija	struktura
jedinica	sustav
jedrilica	tijelo
kult	tvorevina
liječnik	udar
militarac	vlada
militarist	vlast
ministar	vojniki
ministarstvo	vojska
policija	žurnalizam
politika	
postrojba	



**re as E1**

**134 nouns in SrpKor**

adaptacija	industrializacija	konfiguracija	šešeljizacija
afirmacija	instalacija	konsolidacija	sistematizacija
afirmisanje	instaliranje	konstituisanje	socijalizacija
aktiviranje	institucionalizacija	konstrukcija	staljinizacija
aktualizacija	integracija	konstruiranje	standardizacija
alokacija	integrisanje	konstruisanje	stilizacija
animiranje	interpretacija	kontekstualizacija	struktuiranje
aranžiranje	interpretiranje	konzervacija	strukturalizacija
aranžman	intervencija	konzerviranje	strukturiranje
artikulacija	introdukcija	militarizacija	strukturisanje
artikulisane	investiranje	mitologizacija	strukturizacija
autorizacija	inženjering	modelacija	teritorijalizacija
balans	inženjerstvo	modeliranje	teritorijalizovanje
definicija	islamizacija	modelovanje	testiranje
definisanje	izbor	normalizacija	tradicionalizacija
distribucija	izbornost	obdukcija	transkripcija
distribuiranje	izdanje	obrazovanje	transmisija
distributivnost	izdavanje	okupacija	transplantacija
dizajn	izgradnja	operacija	traumatizacija
dizajniranje	izvođenje	organizacija	urbanizacija
edukacija	izvoz	organizovanje	vakcinacija
ekshumacija	judeizacija	osiguranik	vakcinisanje
eksport	kanalizacija	osiguranje	valorizacija
emitovanje	kanalizovanje	osnivanje	valorizovanje
etatizacija	kapitalizacija	otvaranje	vitalizacija
evaluacija	katalogizacija	patrijarhalizacija	vitalizovanje
evidencija	kategorizacija	polarizacija	
federalizacija	klasifikacija	pozicioniranje	
feudalizacija	klasifikovanje	privatizacija	
finansiranje	klerikalizacija	problematiziranje	
formulisanje	kolonizacija	program	
fotografisanje	kombinacija	programiranje	
humanizacija	kombinovanje	projektovanje	
identifikacija	kompenzacija	rafinacija	
identifikovanje	komponovanje	sagledavanje	
indeksiranje	kompozicija	saznanje	

**141 nouns with *re* as E1 in HNK**

adaptacija	emitiranje	kompozicija	populacija
afirmacija	etatizacija	kondicioniranje	popularizacija
afirmiranje	evaluacija	konekcija	pozicija
aktiviranje	evangelizacija	konfiguracija	pozicioniranje
aktivnost	federalizacija	konfiguriranje	privatizacija
aktualizacija	financiranje	konsolidacija	procesiranje
aktualiziranje	habzburgizacija	konstruiranje	profesionalizacija
amaterizacija	identifikacija	konstrukcija	profiliranje
ambulacija	imenovanje	kontrola	program
animacija	industrijalizacija	konverzija	programiranje
animator	instalacija	kristijanizacija	projekt
aranžiranje	instaliranje	kroatizacija	redigiranje
balans	institucionalizacija	kultivacija	reizdanje
balastiranje	integracija	kultiviranje	ruralizacija
balkanizacija	integriranje	kvalificiranje	sistematizacija
bestijalizacija	intepretacija	kvalifikacija	socijalizacija
boljševizacija	interpretacije	licenciranje	srbizacija
bosnizacija	interpretiranje	lokacija	standardizacija
centralizacija	introdukcija	mitologizacija	strukturacija
cirkulacija	investicija	modeliranje	strukturiranje
cirkuliranje	investiranje	monetarizacija	tarifiranje
dajtonizacija	islamizacija	nacionalizacija	tradicionalizacija
definicija	izbor	narativizacija	transmisija
definiranje	izdanje	okupacija	transplantacija
deponiranje	izdavanje	operacija	traumatizacija
destilacija	izgradnja	organizacija	unifikacija
deustašizacija	izvoz	organizator	unitarizacija
dimenzioniranje	kategorizacija	organiziranje	urbanizacija
distribucija	katoliziranje	orijentacija	vakcinacija
distribuiranje	klasificiranje	osiguranje	validacija
dizajn	klasifikacija	osiguravatelj	valorizacija
dizajniranje	klerikalizacija	otkup	vitalizacija
dresiranje	kolonijalizacija	patrijacija	vitaliziranje
educiranje	kombinacija	planiranje	
edukacija	kompiliranje	plasiranje	
eksport	komponiranje	plasman	

## **Appendix C**

In this appendix I have listed a selection of the types of nouns in which I found the selected postposed semiwords and suffixes. I have only listed nouns which certainly or probably have been formed in either Croatian or Serbian. Names of companies, businesses, tv-programmes etc. are not included.

The queried postposed elements were:

### Postposed semiwords/Suffixes

- art
- bus
- teka
- holičar
- holik
- manija
- man
- logija
- log
- ijada
- ator/-itor
- er
- erija
- ant
- izam

**art as E2**

in SrpKor
rat-art

in HNK
-

**ant as E2**

in SrpKor	
cirkusant	praktikant
folirant	prevarant
inkasant	projektant
jurišant	reprezentant
kalkulant	šifrant
kapitulant	šmirant
kolaborant	specijalizant
komplotant	spekulant
kopirant	špekulant
kritikant	trafikant
manifestant	zabušant
markirant	zajebant
muzikant	
narkotrafikant	
politikant	

in HNK	
cirkusant	praktikant
dešifrant	prevarant
folirant	probant
kalkulant	projektant
kapitulant	provokant
kolaborant	reprezentant
kooperant	šifrant
kritikant	šmirant
manifestant	specijalizant
markirant	spekulant
muzikant	špekulant
narkotrafikant	trafikant
politikant	zabušant
poslijediplomant	zafrkant
postdiplomant	zajebant

**ator as E2**

in SrpKor
drmator
gnjavator
grebator
muljator
muvator
snagator
zanimator
zezator
rimator
urlator

in HNK
brbljator
drmator
drpator
gnjavator
grebator
muljator
snagator

In this list all derivations of verbs ending in -irati and nominal verbs ending in '-izacija' as e.g. *evropeizator* (< evropeizirati / evropeizacija) and *tuđmanizator* (tuđmainizirati / tuđmanizacija) are omitted.

**bus as E2**

<b>in SrpKor</b>
bibliobus
konjobus

<b>in HNK</b>
bibliobus

**er as E2**

<b>in SrpKor</b>	
asfalter	masažer
autostoper	montažer
darker	plantažer
fetišer	serviser
finansijer	šminker
grafiter	švercer
hazarder	teniser
kompromiser	tračer

<b>in HNK</b>	
ambalažer	kladitelj-hazarder
aranžer	kompromiser
asfalter	masažer
autobuser	plinoserviser
automasažer	tračer
autostoper	švercer
darker	šminker
djelatnik-serviser	montažer
franšizer	plantažer
grafiter	

**erija as E2**

<b>in SrpKor</b>
donkihoterija
fazanerija
koketerija
korseterija
kruasanterija
partizanerija
pikanterija
piraterija
studenterija

<b>in HNK</b>
cirkuserija
klaunerija
koketerija
kulterija
piraterija

**holik as E2**

in SrpKor
rado <b>h</b> olik

in HNK
rado <b>h</b> olik
čokolado <b>h</b> olik

**holičar as E2**

in SrpKor
blog <b>o</b> holičar
čoko <b>h</b> oličar
kupo <b>h</b> oličar
neto <b>h</b> oličar
rado <b>h</b> oličar

in HNK
dućano <b>h</b> oličar
rado <b>h</b> oličar
šoping <b>o</b> holičar

**ijada as E2**

in SrpKor	
aprilijada	klovnijada
atonijada	kobasicijada
balkanijada	kotličijada
bekrijada	kupusijada
betovenijada	majmunijada
biblijada	motorijada
biciklijada	pasuljijada
bolonjijada	pidžamijada
bombardijada	pitijada
borhesijada	prasicijada
bostanijada	pršutijada
bubijada	rakijada
buregdžijada	rašomonijada
cirkusijada	robinzonijada
čivijada	rolerijada
čvarkijada	roštiljijada
đavolijada	šarlatanijada
džipijada	slaninijada
fijakerijada	šljivijada
folklorijada	sponzorijada
fudbalijada	štrašijada
gitarijada	štrudlijada
gulašijada	tradicijada
gurmanijada	tucanijada
hristijada	univerzijada
kafkijada	vavilonijada
klikerijada	zečijada

in HNK	
amerikanijada	kobasijada
balkanijada	kolačijada
balonijada	kulenijada
bandićijada	kulinijada
biciklijada	kvadrijada
bikijada	lampionijada
bogumilijada	maturijada
brucošijada	motociklijadama
brudetijada	motorijada
cirkusijada	palačinkijada
clintonijada	pametijada
dinosaurijada	panmundžomijada
dionizijada	papirijada
filmijada	pitlijada
fišijada	pivijada
fregolijada	rakijada
ginjoliada	rašomonijada
gitarijada	robinzonijada
golijada	rockijada
humanijada	romobilijada
jamesbondijada	salamijada
jeremićijada	somijada
kartonijada	tamburijada
kestenijada	tuđmanijada
klikerijada	univerzijada
kobasičijada	

***itor as E2***

*-itor* is not, according to my findings, a productive suffix in Serbian and Croatian word-formation.

***izam as E2***

<b>in SrpKor</b>	<b>in HNK</b>
anegdotizam	bandićizam
bajaderizam	banditizam
banditizam	clintonizam
bergsonizam	donkihotizam
biciklizam	druidizam
bušizam	džentlmenizam
fahidiotizam	hadezeizam
familjarizam	hrvatizam
fetišizam	huliganizam
gangsterizam	idiotizam
huliganizam	igmanizam
idiotizam	ikebanizam
klintonizam	jednopartizam
okultizam	lezbijanizam
putinizam	miloševićizam
snobizam	opskurantizam
tragizam	paušalizam
tuđmanizam	pučizam
tupizam	skautizam
višepartizam	snobizam
zombizam	titoizam
	tragizam
	tuđmanizam
	uništizam
	vagabundizam

***log as E2, preceded by -o-***

<b>in SrpKor</b>	<b>in HNK</b>
ćutolog	filmolog
evrolog	fukarolog
filmolog	hamburgerolog
holmsolog	igračkolog
mocartolog	kajkavolog
Njujorkolog	kajolog
rembrantolog	krležolog
romolog	nogometolog
satirolog	parčićolog
šekspirolog	petrarkolog
sovjetolog	petrićolog
tarotolog	shaekespearolog
tramvajolog	televiziolog
turizmolog	vatikanolog
ufolog	

***logija as E2, preceded by -o-***

<b>in SrpKor</b>	<b>in HNK</b>
bratologija	banologija
ciganologija	bubnjologija
citatologija	držićologija
čudovištologija	filmologija
ćuftologija	jajologija
čvorologija	kolumbologija
džojnologija	kremljologija
hristologija	kristologija
imidžologija	krležologija
kaminologija	licologija
kišologija	lupetologija
klozetologija	novitologija
mravologija	papirologija
njegošologija	šatrologija
papirologija	šekspirologija
romologija	šimićologija
srećologija	titologija
stihologija	vinologija
zaverologija	
zmajologija	



**man as E2, preceded by -o-**

in SrpKor	in HNK
brankoman	kavoman
čokoladoman	srkoman
glumoman	grafoman
grafitoman	hašoman
igloman	krležoman
lotoman	stališoman
tabletoman	
srboman	

**manija as E2, preceded by -o- OR a hyphen**

in SrpKor		in HNK	
Ajversonomanija	mađaromanija	apartomanija	romobilomanija
amfetomanija	madonomanija	apsintomanija	sajmomanija
anglomanija	mobilomanija	barkomanija	sektomanija
barbituromanija	Mocart-manija	bombonomanija	šerfezomanija
baštomanija	mocartomanija	cellomanija/ čelomanija	škegromanija
bekamomanija	obamomanija	cestomanija	slavenomanija
Bekam-manija	pogibeljomanija	ćiromanija	špijunomanija
bitlsomanija	pogledomanija	citatomanija	srbomanija
brankomanija	poteromanija	egiptomanija	strancomanija
brendomanija	pračetomanija	festivalomanija	stranomanija
citatomaniya	restoranomanija	festivalomanija	strendžeromanija
dijanomanija	retromanija	fotografomanija	tamburomanija
diskomanija	Ronaldinjomanija	fridomanija	udrugomanija
dopingomanija	sajmomanija	Gavranomanija	vampiromanija
džeksomanija	Šekspiromanija	goranomanija	Zidanomanija
džezomanija	silikomanija	janicomaniya	Zoomanija
fejsbukomanija	sisomanija	kahlomanija	zvjezdomanija
festivalomanija	sloteromanija	kvadratomanija	
festomanija	sovjetomanija	kvizomanija	
filmomanija	špijunomanija	Ljubomanija	
flašomanija	spomenikomanija	Matrixomanija	
gadžetomanija	stentomanija	mobitelomanija	
Jugomanija	sudomanija	peđomanija	
kioskomanija	Telefonomanija	pikadomanija	
krizomanija	televizomanija	pivomanija	
kvizomanija	titomanija	pogibeljomanija	
lažomanija	Van Gog-manija	Potteromanija/ poteromanija	
lotomanija	zvezdomanija		

*teka as E2, preceded by -o-*

<b>in SrpKor</b>	<b>in HNK</b>
biljoteka	filmoteka
eroteka	igroteka
filmoteka	vinoteka
folkoteka	mikroteka
mikrofilmoteka	folkoteka
programoteka	
restoteka	
vicoteka	
videoteka	
vinoteka	
željoteka	

## **Appendix D**

This appendix contains the detailed description of how the nouns (1), nominal compounds (2), semiwords, affixes (3) and indefinite markers (4) were queried in SrpKor and HNK.

All the queried nouns are introduced in Table 4.4.1-4.4.6.

All the results of the queries of elements in compound-like constructions: nominal compounds, semiwords and affixes are listed in appendix A, B and C.

## 1. Querying nouns

Due to different possible spellings of the replicated nouns and sometimes of the Serbian and Croatian equivalents, searches were made for sometimes up to six different lemmas in order to determine the number of instances of one replicated noun. Thus, in order to obtain the absolute number of instances of the lemma [time-out], searches were conducted for the following six lemmas in both corpora: {time-out, time out, timeout, tajm-aut, tajm aut, tajm aut}. Because of the micro-dialectal difference between Croatian (Ijekavian reflex of the Common Slavic *jat*) and Serbian (Ekavian and Ijekavian reflex of the Common Slavic *jat*) nouns containing *jat* (e.g. *tjelovježba*) were queried in two versions: *telovežba* and *tjelovježba*. Differences in adaptation of replicated matter such as Serbian: *funkcioner*, *regrutovanje* / Croatian: *funkcionar*, *regrutiranje* ('official, recruitment') were also included in the queries but in the following tables, one lemma e.g. *tajm aut* or *funkcioner* is used as a representative, covering all versions of spelling and adaptation of that lemma. The replicas fall into the category of 'adaptees' as they have all been adapted morphologically (transmorphemized) to Serbian and/or Croatian as they are assigned a gender (and nominal declension), and in most instances also a domestic suffix (-ija, -a, -Ø) to accommodate the gender assignment. They have also been adapted phonologically (transphonemized), some to the extent that word final consonant clusters and vowel clusters are evaded (e.g. *koktel* /'kɔktɛl/ < cocktail), while others retain this feature (e.g. *bodibilding* /'bɔdibilding/ < bodybuilding, *tajm aut* /'tɔjm aut/ or /'tɔjm' aut/ < time-out).

The searched corpora of running text are lemmatised, so it is possible to search for lemmas to determine the total amount of wordforms of the same lemma (lexeme). However, querying the corpora, I found that not all word-forms are lemmatised, which became evident when a search of a lemma (e.g. *period*) only resulted in concordances with the wordform *period* and not all the possible declined wordforms of the noun. In these instances, I subsequently searched for all possible wordforms of the queried lexeme in question. Concretely, instead of making the query: [lemma = "period"], I made the query [wordform = "period(a|u|om|i|e|ima)?"] and thus gained access to all instances of: *period*, *perioda*, *periodom*, *periodi*, *periode*, *periodima*, which are all possible forms (7 cases and in singular and plural) of the noun *period* in either HNK or SrpKor.

## 2. Querying nominal compounds

As was the case with the queried nouns, the corpora contain different spellings of the two nouns combined in compounds. Therefore, several queries were made of each noun in both corpora so that the results would comprise all compounds with the queried E1 or E2 of [N<sub>UNDECL</sub>N]<sub>N</sub>-compounds.

In SrpKor the queries were made by choosing the option “napredna pretraga” (advanced search), and by entering the following sequences in the search field I obtained a sample of all compounds in which a particular E1 was present, here exemplified with *kontakt*:

Query 1a: [word = "(k|K)ontakt"] [pos = "N"]

Query 2a: [word = "(k|K)ontakt "] [word = "-"] [pos = "N"]

Through query 1a all instances of the token *Kontakt* or *kontakt* followed by any noun was found, e.g.:

*“za Jevreje postoje **kontakt adrese**”* (SrpKor)  
(‘...for Jews **contact addresses** exist’)

Through query 2a all instances of *Kontakt* or *kontakt* followed by a hyphen, which is followed by any noun was found, e.g.:

*“..okolnosti pod kojima je došlo do oštećenja novčanice , mesto i datum predaje novčanice banci , odnosno filijali , **kontakt - telefon** , adresu , ime , prezime i jedinstveni matični broj građana lica koje novčanicu podnosi na zamenu”* (SrpKor)

(‘..the circumstances under which the note was damaged, place and date when the note is handed in at the bank or branch-office, **contact telephone number**, address, name, surname, and ID-number of the person who is applying for an exchange of the note’)

To obtain a sample of all compounds in which the E2 was a specific noun, the SrpKor was queried in the following way, here exemplified by *šou*<sup>306</sup> (‘show’):

---

<sup>306</sup> As it is not possible to search for letters with diacritics the digraphs of Serbian in SrpKor, those letters are searched for in the following way: ć/Ć - cx/Cx, č/Č - cy/Cy, , đ/Đ - dx/Dx, dž/Dž - dx/Dx, š/Š - sx/Sx, ž/Ž - zx/Zx, lj/Lj - lx/Lx, nj/Nj - nx/Nx.

Query 3a: [pos = "N"] [lemma = "sxou"]

Query 4a: [pos = "N"] [word="-"] [lemma = "sxou"]

Through query 3a all instances of any noun followed by any form of *šou* was found, e.g.:

*"Tročasovni televizijski **gala šou** iz Osla, u dvorani "Spektrum" pred oko 7. 000 gledalaca"* (SrpKor)

('The three hour long **gala show** from Oslo in the "Spektrum" hall in front of around 7,000 viewers').

Through query 4a any noun combined via a hyphen to any form of *šou* was found, e.g.:

*"..dok će specijalni **web - šou** voditi američka glumica i, odnedavno, pevačica Džulijet Luis"* (SrpKor)

('..whereas the special **web show** will be led by the American actress and, of recent, singer Juliet Lewis.')

In HNK, the queries for [NUNDECLN]<sub>N</sub>-compounds containing a specific E1 was conducted in a manner similar to the one applied in SrpKor with the exception that hyphens as a rule are not tokenised in the HNK, even though I detected some instances in which they were. Accordingly, the following queries (here exemplified with *internet*) were conducted by opting for a search using the corpus querying language (CQL) and by entering the following sequences in the search field:

Query 1b: [word = "(i|I)nternet"] [msd="N.\*"]

Query 2ba: [word="(i|I)nternet -.\*" & msd="N.\*"]

Query 2bb: [word = "(i|I)nternet"] [word = "-"] [msd="N.\*"]

Through query 1b all instances of *internet* or *Internet* followed by any noun were found, e.g.:

*"Ova informacija može se provjeriti na siteu Ministarstva vanjskih poslova na **Internet adresi** mvp.hr."* (HNK)

('This information may be checked at the homepage of The Ministry of Foreign Affairs at the **Internet address** mvp.hr.')

Through query 2ba and 2bb all instances of the noun *internet* followed by a hyphen and then any noun was found, e.g:

*“Najčešće se radi o tvrtkama čiju veću **internet-angažiranost** nalažu inozemni poslovni partneri”* (HNK)

(‘It most often relates to companies whose larger **internet-involve-ment** is demanded by foreign business partners’)

*“**Internet - usluga** ugovorena između imatelja transakcijskog računa i Banke kojom je imatelju transakcijskog računa omogućeno zadavanje naloga”* (HNK)

(‘An **Internet service** agreed upon by the account owner and the bank through which the account owner may give a transfer/payment order’)

To obtain a sample of all compounds in which the E2 was a specific noun, the HNK was queried in the following way, here exemplified by *zona*

Query 3b: [msd = "N(c|p)(m|n|f)sn.\*"] [lemma = "zona"]

Query 4ba: [msd="N.\*" & word=".\*-zon(a|u|e|i|om|o|ama) "]

Query 4bb: [msd = "N(c|p)(m|n|f)sn.\*"] [word = "-"] [lemma = "zona"]

Through query 3b all instances of the noun *zona* preceded by a noun (N) which is either a proper or common noun (c|p) and is masculine, neuter or feminine (m|n|f) in the singular (s), nominative (n) form. Because the HNK is automatically tagged with morphosyntactic description, it is possible in the query to eliminate all other forms of nouns, i.e. nouns in the plural and declined nouns (not in nominative), e.g.:

*“Među zemljama istočne Europe više nema **tampon zone**, a kontrola je zajednička”* (HNK)

(There is no longer a **buffer zone** between the countries of Eastern Europe, and there is joint [border] control)

Through query 4ba all nouns (N) ending in all possible forms of *zona* preceded by a string of letters (\*) beginning with any letter (a-z) followed by a hyphen (-) are found, e.g.:

*“I on je protiv najnovijih nelogičnih prostornih rješenja prema kojima je njegova kuća u **vikend-zoni**.”* (HNK)

(He is also against the illogical areal decisions according to which his house is situated in the **weekend-zone** (i.e. recreational area))

Through query 4bb all nouns in the nominative (undeclined) which precede a hyphen which precedes the lemma *zona* are found, e.g.:

*“Tenkovi JNA su u ljeto 1991. ušli u Okučane , oponašajući **tampon - zonu** između srpskih pobunjenika i hrvatskih policijskih snaga” (HNK)*  
(‘In the summer of 1991, tanks from The Yugoslav People’s Army entered Okučani and pretended to be a **buffer zone** between the Serbian rebels and the Croatian police forces.’)

### **Auto**

In order to find all the possible types of E2s with which *auto* combines in compounds, an additional query was added, because compounds with *auto* are often written as one word, e.g. *autoput* (‘motorway’), *autoguma* (‘car tyre’). The queries conducted were:

Query 5a) In SrpKor: [word = "(A|a)uto.\*" & pos = "N"]

Query 5b) In HNK: [word="(A|a)uto.\*" & msd="N.\*"]

Through these queries all nouns (N) beginning with *auto* or *Auto* were found. Subsequently I eliminated all irrelevant results in which *auto* has the meaning ‘self’ and not automobile.

In compounds where the E1 ends in a –o, as does *auto*, the –o is, according to Piper and Klajn (2013: 249) to be considered a linking element. , which is, I assume, the reason why a rather large number of the different compounds are found in the corpora written as one word (67 out of 151 in SrpKor and 125 out of 134 in HNK). In other words, the word-final –o in *auto* increases the likelihood that compounds are formed with the noun *auto* as E1.



### 3. Querying semiwords and affixes

In SrpKor the queries for preposed semiwords and prefixes were conducted in a manner similar to how nominal compounds were queried, i.e. by opting for “Napredna pretraga” (advanced search) and by entering the following sequences in the search field, exemplified here by the preposed semiword “*eko*”.

Query 6a: [word = "(E|e)ko"] [pos = "N"]

Query 7a: [word = "(E|e)ko"] [word = "-"] [pos = "N"]

Query 8a: [word = "(E|e)ko.\*" & pos = "N"]

In HNK I also queried for preposed semiwords and prefixes in a way resembling the way I queried for nominal compounds. Opting for a search using the corpus querying language (CQL) I conducted my search by entering the following sequences in the search field:

Query 6b: [word = "(E|e)ko"[msd="N.\*"]

Query 7ba:[lemma="(E|e)ko-[a-z]\*" & msd="N.\*"]

Query 7bb: [word = "(E|e)ko"] [ word "-"] [msd="N.\*"]

Query 7bc: [word = "(E|e)ko-"] [msd="N.\*"]

Query 8b: [lemma="(E|e)ko.[a-z]\*" & msd="N.\*"]

Query 6a and 6b result in all instances where the token *eko* precedes a noun, in a sequence of two tokens unhyphenated, e.g. “*eko selo*”, “*eko selu*”, “*eko park*”, “*Eko parkom*”, etc.

Queries 7a and 7b(a-c) result in all instances where *eko* precedes a noun with which it is combined via a hyphen, whether it is written (a) *eko-selo*, (b) *eko - selo* or (c) *eko- selo*.

Queries 8a and 8b result in all instances where the string of letters *eko* is the initial string of letters in a noun, by which all instances of nouns, including properly hyphenated ones (i.e. without spacing) are found, e.g. *ekoselo*, *eko-selo*, *ekonomija*, *ekolog*, etc.

Query 7ba in HNK is thus made redundant, unless I want to know the exact number of hyphenated instances of nouns with *eko* as the initial part.

To obtain a sample of all nouns ending in a specific semiword or suffix, the SrpKor was queried in the following way, here exemplified by *-log*:

Query 9a: [word="\*.lo(g|ga|gu|gom|zxe|zi|ge|zima)"]

Query 10a: [word=".\*"][word="lo(g|ga|gu|gom|zxe|zi|ge|zima)"]

Query 11a: [word=".\*"] [word = "-"] [word = "lo(g|ga|gu|gom|zxe|zi|ge|zima)"]

The HNK was queried in a similar way:

Query 9b: [word="\*.lo(g|ga|gu|gom|zi|ge|zima|že)"]

Query 10b: [word=".\*"][word="lo(g|ga|gu|gom|zi|ge|zima|že)"]

Queries 9a and 9b result in all wordforms which end in *log* or any of the forms *-log* may take in the declension of masculine nouns ending in *log*. Query 9b also resulted in all instances where *log* was preceded by a hyphen in HNK, whereas query 11a was used for this purpose in SrpKor. Finally, queries 10a and 10b resulted in all instances where there was a space between any token and the queried form.

In most instances, no results were found through queries 10ab and 11a, as the majority of the postposed semiwords and presumably all of the suffixes are never combined with a preceding lexical stem or lexeme through hyphenation or spacing.

However, when querying the postposed semiwords *-bus*, *-art* and *m-anija*, some wordforms with both hyphenation and spacing were found, e.g. *mocart-manija* (Mozart-mania), *rat-art* (war art), *grafiti art* (graffiti art), *mini bus* (mini bus).

As I only searched for nouns, all the nouns found via the searches with preposed element (semiwords and prefixes) consist of [semiword/prefix + noun] or [semiword+semiword]. The searches for postposed elements (semiwords and derivational suffixes) may however have lexical stems from different word classes, as it is the postposed element that defines the word class of the formed word. So, the nouns found in these queries are best described as [lexical stem + semiword]<sub>N</sub> or [lexical stem + nominal suffix]<sub>N</sub>

#### 4. Querying *jedan* and *neki*

In SrpKor the queries of the 8 NPs in the singular were carried out by using the option “jednostavna pretraga” (simple search) in which the possible forms of each NP was queried, here exemplified by *jedan prijatelj*:

Query 12a: (|j|)edn(oga|og|om|ome|omu|im|ime) prijatelj(a|u|em|om)

Query 13a: (|j|)jedan prijatelj

In HNK, the queries were carried out in a similar manner by using the option “Phrase”:

Query 12b: (|j|)edn(oga|og|om|ome|omu|im|ime) prijatelj(a|u|em)

Query 13b: (|j|)jedan prijatelj

Queries 12a-b and 13a-b result in concordances with all the instances of *jedan prijatelj* and *Jedan prijatelj* in the nominative, accusative, genitive, dative, instrumental and locative cases in Croatian and Serbian.

The vocative case was omitted as *jedan* and *neki* are not expected to be used as articles in the vocative. *jedan* may be used as a postposed determiner of a noun but is then not article-like. *jedan*’s meaning is under such circumstances augmentative, cf. *Lopove jedan!* (‘You thief you!’ / ‘What a thief you are!’)

As with the other searches, the differences between the Ekavian and Ijekavian versions as well as differences in pronunciation and spelling between Standard Serbian and Standard Croatian were considered. For instance: *djevojka* (Ijekavian) and *stol* (Croatian preference for wordfinal /l/ were also queried as *devojka* (Ekavian) and *sto*.

#### Querying for generic use of *jedan* and *neki*

Additional queries were carried in order confirm that *jedan* and *neki* are not in used to mark generic reference. This was done by entering the following queries:

Query 14a [word = "posta(o|la|lo|li|le)"] [lemma = "jedan|neki"] [pos = "N"]

Query 14b [word = "posta(o|la|lo|li|le)"] [lemma = "jedan|neki"]  
[msd="N.\*"]

Queries 14a and 14b result in concordances of POSTAO+[NEKI+N]<sub>NP</sub> and "POSTAO+[JEDAN+N]<sub>NP</sub>".

The wordforms *postao*, *postala*, *postalo*, *postali* and *postale* represent all the possible forms of the perfect participle active of the verb *postati* ('to become'). The perfect participle active act as the main verb in VPs which constitute the following verbal tenses and moods, her exemplified by the participle in the masculine, singular form, *postao*:

The perfect:	<i>On je postao</i>	- 'He became/ has become'
The past perfect:	<i>On je bio postao</i>	- 'He had become'
The future perfect:	<i>Kad on bude postao</i>	- 'When he has become'
The subjunctive I:	<i>On bi postao</i>	- 'He would (like to) become'
		'He would have become'
The subjunctive II:	<i>On bi bio postao</i>	- 'He would have (had) become'

## Bibliography

- Alexander, Ronelle. 2006. *Bosnian, Croatian, Serbian, a Grammar, with Sociolinguistic Commentary*. University of Wisconsin Press
- Ammon, Ulrich. 1989. *Status and Function of Languages and Language Varieties*. Berlin, Boston: De Gruyter
- Andersen, Henning. 2001. *Actualization, Linguistic Change in Progress*. Current Issues in Linguistic Theory v. 219. John Benjamins PubCo
- . 2008. 'Grammaticalization in a Speaker-Oriented Theory of Change'. *Grammatical Change and Linguistic Theory* 2008: 11–44
- Anić, Vladimir. 1990. 'Prilog razgovoru o planiranju jezika u Jugoslaviji'. In *Jezička politika i planiranje u Jugoslaviji*, 101–2. Novi Sad, Serbia: Institut za južnoslovenske jezike, Filozofski fakultet u Novom Sadu
- Annamalai, E. 1979. 'Movement for Linguistic Purism. The Case of Tamil'. In *Language Movements in India*, edited by E. Annamalai, 35–39. Mysore: Central Institute of Indian Languages
- . 1989. 'The Linguistic and Social Dimensions of Purism'. In *The Politics of Language Purism*, edited by Björn H. Jernudd and Michael J. Shapiro, 225–32. Berlin, Boston: De Gruyter Mouton.
- Auburger, Leopold. 1999. *Die kroatische Sprache und der Serbokroatismus*. Ulm: Gerhard Hess Verlag
- Auer, Peter. 2005. 'Europe's Sociolinguistic Unity, or: A Typology of European Dialect/standard Constellations'. In *Perspectives on Variation, Sociolinguistic, Historical, Comparative*, edited by Nicole Delbecque, Johan van der Auwera, and Dirk Geeraerts. Mouton de Gruyter
- Auwera, Johan van der. 2002. 'More Thoughts on Degrammaticalization'. In *New Reflections on Grammaticalization*, edited by Ilse Wischer and Gabriele Diewald, 19–29. Philadelphia, USA: John Benjamins Publishing Company
- Babić, Stjepan. 1971. 'Tvorba imenica na -ač'. *Radovi Zavoda za slavensku filologiju* 12: 13–20
- . 1986. *Tvorba riječi u hrvatskom književnom jeziku*. Djela Jugoslavenske akademije znanosti i umjetnosti, Razred za filologiju 62. Zagreb: Jugoslavenska akademija znanosti i umjetnosti, Globus
- . 2009. 'Die kroatische Literatursprache gestern und heute'. In *Jezični varijeteti i nacionalni identiteti, prilozi proučavanju standardnih jezika utemeljenih na štokavštini*, edited by Lada Badurina, Ivo Pranjković, and Josip Silić, 169–77. Zagreb: Disput
- Babić, Stjepan, Božidar Finka, and Milan Moguš. 1990. *Hrvatski pravopis*. 4. Zagreb: Školska knjiga
- Babić, Stjepan, and Sanda Ham. 2005. 'Pravopisni rat'. *Jezik, časopis za kulturu hrvatskoga književnog jezika* 52 (1–4): 22–30–65–109–148
- Babić, Stjepan, Sanda Ham, and Milan Moguš. 2008. *Hrvatski školski pravopis - usklađen sa zaključcima Vijeća za normu hrvatskoga standardnog jezika*. Zagreb: Školska knjiga

- Babić, Stjepan, and Milan Moguš. 2011. *Hrvatski pravopis - usklađen sa zaključcima Vijeća za normu hrvatskoga standardnog jezika*. 2. Zagreb: Školska knjiga
- Baotić, Josip. 2005. 'Književnojezička politika 1970-1990 - Borba za zajedništvo i ravnopravnost'. In *Jezik u Bosni i Hercegovini*, 435–77. Oslo: Institut za istočnoevropske i orijentalne studije
- Barić, Eugenija. 1980. *Imeničke složenice neprefiksalne i nesufiksalne tvorbe*. Znanstvena biblioteka Hrvatskog filološkog društva ; 11. Sveučilišna naklada Liber
- Barić, Eugenija et al. 1999. *Hrvatski jezični savjetnik*. Edited by Lana Hudeček, Milica Mihaljević, and Luka Vukojević. Zagreb: Institut za hrvatski jezik i jezikoslovlje - Pergamena - Školske novine
- Barić, Eugenija, Mijo Lončarić, Dragica Malić, Slavko Pavešić, Mirko Peti, Vesna Zečević, and Marija Znika. 1995. *Hrvatska gramatika*. Zagreb: Školska Knjiga
- Barjamović Orahovski, Miro. 2006. *Lingvistisk terminologi - Serbokroatisk-dansk glosar*. Vol. 1. 2 vols. Arbejdspapirer. Århus: Slavisk Institut
- Bašić, Nataša. 2012. 'Hrvatski jezik i jezikoslovlje danas: normativni prijevori i politička osporavanja'. Edited by Sanda Ham. *Jezik, časopis za kulturu hrvatskoga književnog jezika* 59 (2): 60–71
- . 2014. 'U povodu objave Hrvatskoga pravopisa Instituta za hrvatski jezik i jezikoslovlje'. *Jezik: Periodical for the Culture of the Standard Croatian Language* 61 (December): 183–89
- Bauer, Laurie. 1988. *Introducing Linguistic Morphology*. Edinburgh University Press
- . 1998. 'Is There a Class of Neoclassical Compounds, and If so Is It Productive?'. *Linguistics* 36 (355): 403–422
- . 2002. 'Inferring Variation and Change from Public Corpora'. In *The Handbook of Language Variation and Change* -, edited by J.K. Chambers, Peter Trudgill, and Natalie Schilling, 97–114. UK: Blackwell Publishing
- Bekavac, Božo, Sanja Seljan, and Ivana Simeon. 2008. 'Corpus-Based Comparison of Contemporary Croatian, Serbian and Bosnian'. In *Proceedings*, edited by Marko Tadić, Mila Dimitrova-Vulchanova, and Svetla Koeva, 33–39. Zagreb, Croatia: Croatian Language Technologies Society - Faculty of Humanities and Social Sciences
- Belić, Aleksander. 1923. *Pravopis srpskohrvatskog književnog jezika*
- Benson, Morton. 1996. *SerboCroatian-English Dictionary*. Cambridge University Press
- Blommaert, Jan, Sirpa Leppänen, Päivi Pahta, and Tiina Räisänen, eds. 2012. *Dangerous Multilingualism*. Palgrave Macmillan.
- 'Bolje je hrvatski'. 2015. *Bolje je hrvatski*. <http://bolje.hr/>.
- Boranić, Dragutin. 1921. *Pravopis hrvatskoga ili srpskoga jezika*. 1. Zagreb: Naklada Ćirilo-Methodske nakladne knjižare
- Brborić, Branislav. 1996. 'Odnos prema tuđicama: S merom ali bez averzije'. In *O leksičkim pozajmljenicama*, edited by Judita Plankoš, 27–

51. Subotica - Beograd: Institut za srpski jezik SANU, Gradska biblioteka Subotica
- . 2001. *S jezika na jezik*. Beograd: Centar za primenjenu lingvistiku: 'Prometej'
- . 2011. 'Standardni jezik i jezički standard'. In *Srpski jezički priručnik*, 5th edition, 11–36. Beograd: Beogradska knjiga
- Brborić, Branislav, Jovan Vuksanović, and Radojko Gačević. 2006. *Srpski jezik u normativnom ogledalu : 50 odluka Odbora za standardizaciju srpskog jezika*. Beograd : Beogradska knjiga,
- Brborić, Veljko. 2008. 'Srpski pravopis u dvadesetom veku'. *Zbornik Matice srpske za slavistiku* 73
- . 2013. 'Srpski pravopis na prelazu iz XIX u XX vek u svetlu današnjih rešenja'. *Intralingvistički i ekstralingvistički činiooci u formiranju i razvoju srpskog standardnog jezika*, Naučni sastanak slavista u Vukove dane, 42/1: 173–85
- Brinton, Laurel J., and Elizabeth C. Traugott. 2005. *Lexicalization and Language Change*. Cambridge University Press
- Brozović, Dalibor. 1970. *Standardni jezik; teorija, usporedbe, geneza, povijest, suvremena zbilja*. Zagreb: Matica hrvatska
- . 1978. 'Hrvatski jezik, njegovo mjesto unutar južnoslavenskih i drugih slavenskih jezika, njegove povijesne mijene kao jezika hrvatske književnosti'. In *Hrvatska književnost u evropskom kontekstu*, 9–83. Zagreb: Zavod za znanost o književnosti Filozofskog fakulteta Sveučilišta u Zagrebu
- . 1985. 'O funkcioniranju jezika u jugoslavenskoj federaciji'. *Književni jezik* 14 (2): 65–85
- . 1990. 'O općoj tipologiji jezičnih politika i o načelima i uvjetovanostima jezične politike u jugoslavenskoj zajednici'. In *Jezička politika i planiranje jezika u Jugoslaviji*, 7–22. Novi Sad: Institut za južnoslovenske jezike, Filozofski fakultet u Novom Sadu
- . 1998. 'Povijesna podloga i jezičnopolitičke i sociolingvističke okolnosti'. In *Hrvatski jezik*, edited by Mijo Lončarić, 3–33. Opole: Uniwersytet Opolski, Instytut Filologii Polskiej
- . 2003. 'O pravopisima općenito, a o hrvatskima posebno'. Edited by Nataša Bašić. *Hrvatski pravopis 1892-2002*, Dometi, 1–4/2002: 13–23
- Brozović, Dalibor, and Pavle Ivić. 1988. *Jezik, srpskohrvatski/hrvatskosrpski, hrvatski ili srpski*. Zagreb: Jugoslavenski leksikografski zavod 'Miroslav Krleža'
- Bugarski, Ranko. 1996. 'Strane reči danas: Pojam, upotreba, stavovi'. In *O leksičkim pozajmljenicama*, edited by Judita Plankoš, 17–25. Subotica - Beograd: Institut za srpski jezik SANU, Gradska biblioteka Subotica

- . 2004. 'Language and Boundaries in the Yugoslav Context'. In *Language, Discourse, and Borders in the Yugoslav Successor States*, edited by Birgitta Busch and Helen Kelly-Holmes. Multilingual Matters Ltd
- . 2009. *Evropa u jeziku*. Asteroidi edition. Beograd: XX vek, Krug
- . 2012. 'Language, Identity and Borders in the Former Serbo-Croatian Area'. *Journal of Multilingual and Multicultural Development* 33 (3): 219–35
- Bujas, Željko. 2008. *Veliki englesko-hrvatski rječnik*. 4. Zagreb: Nakladni zavod Globus
- Busch, Brigitta. 2010. 'New National Languages in Eastern Europe'. In *The Handbook of Language and Globalization*, edited by Nikolas Coupland, 182–200. Wiley-Blackwell
- CESAR, Central and Southeast European Resources. 2013. *CIP-Pilot Actions Project No. 271022 - Deliverable D3.3 (Part A)*. Version No. 1.0. CESAR - Central and Southeast European Resources.
- Chambers, J.K., Peter Trudgill, and Natalie Schilling, eds. 2002. *The Handbook of Language Variation and Change* -. UK: Blackwell Publishing
- Clyne, Michael. 1997. 'Epilogue'. In *Undoing and Redoing Corpus Planning*, edited by Michael Clyne, 477–502. Berlin: Mouton de Gruyter
- Coca-Cola HBC Hrvatska. 2016. 'Jezični priručnik Coca-Cole HBC Hrvatska'. Accessed January 29. <http://www.prirucnik.hr/index.html>.
- Coseriu, Eugenio. 1974. *Synchronie, Diachronie und Geschichte, das Problem des Sprachwandels*. Vol. 2. Internationale Bibliothek für allgemeine Linguistik. München: Fink
- Crystal, David. 2008. *A Dictionary of Linguistics and Phonetics*. 6. ed. The Language Library. Blackwell Pub
- Czerwiński, Maciej. 2004. 'Language Planning and Lexical Networks in Croatian Media'. *Net Culture Science / Netz Kultur Wissenschaft*. <https://www.academia.edu/3002537/>
- Čupić, Željko. 1996. 'Strana leksika u rubrici "kulturni život"'. In *O leksičkim pozajmljenicama*, edited by Judita Plankoš, 27–51. Subotica - Beograd: Institut za srpski jezik SANU, Gradska biblioteka Subotica
- Ćirilov, Jovan. 1982. *Rečnik novih reči*. Beograd: Narodna knjiga
- Ćorić, Božo. 2008. *Tvorba imenica u srpskom jeziku*. Biblioteka Književnost i jezik 24. Beograd: Društvo za srpski jezik i književnost Srbije
- Ćupić, Drago, Egon Fekete, and Bogdan Terzić. 2002. *Slovo o jeziku - Jezički poučnik*. Edited by Miloš Jevtić. Vol. 2. 2 vols. Belgrade: Partenon
- Dešić, Milorad. 2009. 'Jezik srpskih i hrvatskih medija - leksičko-semantički aspekt'. In *Die Unterschiede zwischen dem Bosnischen/Bosnia-*



- kischen, Kroatischen und Serbische - Lexik - Wortbildung - Phraseologie*, edited by Branko Tošović, 2:115–21. Graz: LIT Verlag Münster
- Despalatović, Elinor Murray. 1975. *Ljudevit Gaj and the Illyrian Movement*. Columbia University Press
- Dobek, Tajana. 2013. 'Jezik i prava nacionalnih manjina u hrvatskom obrazovnom sustavu'. MA thesis, University of Zagreb
- Dobrić, Nikola. 2012. 'Savremeni Jezički Korpusi Na Balkanu - Istorijat, Trenutno Stanje I Budućnost'. *Slavistična Revija* 60/2012 (4): 677–92
- Dragičević, Rajna. 2009. 'Inovacije u tvorbi reči u slovenskim jezicima'. *Južnoslovenski filolog* 65: 461–68
- Drljača Margić, Branka. 2009. 'Engleski element u oglašavanju na hrvatskom i srpskom jeziku'. In *Die Unterschiede zwischen dem Bosnischen/Bosniakischen, Kroatischen und Serbische - Lexik - Wortbildung - Phraseologie*, edited by Branko Tošović, 2:263–79. Graz: LIT Verlag Münster
- Društvo za srpski jezik i književnost Srbije. 2015. 'О Друштву за српски језик и књижевност Србије'. Accessed April 21. [http://drustvosj.fil.bg.ac.rs/?page\\_id=420](http://drustvosj.fil.bg.ac.rs/?page_id=420)
- Duckworth, David. 1977. 'Zur terminologischen und systematischen Grundlage'. In *Werner Betz: Sprachliche Interferenz : Festschrift für Werner Betz zum 65. Geburtstag*, edited by Herbert Kolb and Hartmut Laufer, 36–56. Tübingen: Max Niemeyer Verlag
- Edwards, John R. 1985. *Language, Society, and Identity*. Language Library. Oxford: Basil Blackwell
- Encyclopedia Britannica. 2015. 'Prague School'. *Encyclopedia Britannica*. February 4. <http://global.britannica.com>.
- European Council. 2013. *Report of the Committee of Experts on the European Charter for Regional and Minority Languages*. ECRML (2013) 3. Strasbourg
- Fekete, Egon. 1996a. 'Tuđice'. In *Slovo o jeziku - Jezički poučnik*, edited by Miloš Jevtić, 1:149–83. Belgrade: Partenon
- . 1996b. 'Otpor prema tuđicama'. In *O leksičkim pozajmljenicama*, edited by Judita Plankoš, 27–51. Subotica - Beograd: Institut za srpski jezik SANU, Gradska biblioteka Subotica
- Field, Fredric. 2002. *Linguistic Borrowing in Bilingual Contexts*. Amsterdam/Philadelphia: John Benjamins Publishing Company
- Filipović, Rudolf. 1990. *Anglicizmi u hrvatskom ili srpskom jeziku, porijeklo - razvoj - značenje*. Djela Jugoslavenske akademije znanosti i umjetnosti 70. Jugoslavenska akademija znanosti i umjetnosti
- . 2004. 'Croatian'. In *English in Europe*, edited by Manfred Görlach, 231–40. Oxford: Oxford University Press
- Fishman, Joshua A. 1975. *Sociolinguistics, a Brief Introduction*. 4. printing. Newbury House Language Series. Newbury House

- Gammelgaard, Karen. 2004. 'Pragerskolen og skriftlighedssproget - et tjekkisk perspektiv'. In *Skriftkultur*, edited by Kjell Lars Berge, Trine Gedde Dahl, and Stephen Walton, 59–71. Oslo: Norsk sakprosa og NFF
- Goebl et al. 1997a. *Kontaktlinguistik*. Berlin, New York: Walter de Gruyter
- . 1997b. *Kontaktlinguistik*. Vol. 2. Berlin, New York: Walter de Gruyter
- Görlach, Manfred, ed. 2004. *English in Europe*. Oxford: Oxford University Press
- Gottlieb, Henrik. 2006. 'Linguistic Influence'. Edited by Kirsten Malmkjær. *Encyclopedia of Language and Linguistics*. Oxford: Elsevier Ltd
- Granić, Jagoda. 2009. 'Purizam ili o(d)brana autohtone bosanske/bošnjačke, hrvatske i srpske leksičke norme'. In *Die Unterschiede zwischen dem Bosnischen/Bosniakischen, Kroatischen und Serbische - Lexik - Wortbildung - Phraseologie*, edited by Branko Tošović, 2:65–82. Graz: LIT Verlag Münster
- Grčević, Mario. 2002. 'Some Remarks on Recent Lexical Changes in the Croatian Language'. Edited by Radovan Lučić. *Lexical Norm and National Language. Lexicography and Language Policy in South-Slavic Languages after 1989*, 150–63
- Greenberg, Robert D. 2008. *Language and Identity in the Balkans, Serbo-Croatian and Its Disintegration*. Updated ed. Oxford University Press
- Gröschel, Bernhard. 2009. *Das Serbokroatische zwischen Linguistik und Politik : Mit einer Bibliographie zum postjugoslavischen Sprachenstreit*. München
- Guberina, Petar, and Kruno Krstić. 1977. *Razlike između hrvatskoga i srpskoga književnog jezika*
- Haarmann, Harald. 1989. 'Functional Aspects of Language Varieties - A Theoretical-Methodological Approach'. In *Status and Function of Languages and Language Varieties*, 153–93
- . 1990. 'Language Planning in the Light of a General Theory of Language: A Methodological Framework'. *International Journal of the Sociology of Language* 86 (1): 103–126
- Ham, Sanda. 2007. 'Najbolje nove hrvatske riječi'. Edited by Sanda Ham. *Jezik, časopis za kulturu hrvatskoga književnog jezika* 54 (3): 119–20
- . 2008. 'Dodijeljene Šreterove nagrade za najbolju novu hrvatsku riječ u 2007.'. Edited by Sanda Ham. *Jezik, časopis za kulturu hrvatskoga književnog jezika* 55 (3): 107–9
- . 2009. 'Dodijeljene nagrade za najbolju novu hrvatsku riječ u 2008.'. Edited by Sanda Ham. *Jezik, časopis za kulturu hrvatskoga književnog jezika* 57 (2): 74
- . 2010. 'Jezični natječaj u 2009.'. Edited by Sanda Ham. *Jezik, časopis za kulturu hrvatskoga književnog jezika* 57 (3): 114–15

- . 2011a. 'Dodijeljene Šreterove nagrade za novu hrvatsku riječ u 2010. – ispraznica, osjećajnik, parkomat'. Edited by Sanda Ham. *Jezik, časopis za kulturu hrvatskoga književnog jezika* 58 (2): 73–78
- . 2011b. 'O jeziku poslovne komunikacije u Coca-Coli'. *Jezik: časopis za kulturu hrvatskoga književnog jezika* 58 (5): 193–95
- Hansen, Erik. 1997. *Dæmonernes port, støttemateriale til undervisningen i nydansk grammatik*. 4th ed. Copenhagen: Erik Hansen og Hans Reitzels Forlag A/S
- Harris, Alice. 2003. 'Cross-Linguistic Perspectives on Syntactic Change'. In *The Handbook of Historical Linguistics*, edited by Brian D. Joseph and Richard D. Janda, 529–51. Blackwell Handbooks in Linguistics. Blackwell Pub
- Haspelmath, Martin. 2009. 'Lexical Borrowings: Concepts and Issues'. In *Loanwords in the World's Languages*, 35–54
- Hauge, Kjetil Rå. 1977. *Indefinite and Generic NP's in Bulgarian and the Question of the Indefinite Article*. Oslo
- Haugen, Einar. 1950. 'The Analysis of Linguistic Borrowing'. *Language* 26 (2): 210–31
- Haugen, Einar. 1966. *Language Conflict and Language Planning, The Case of Modern Norwegian*. Cambridge, Mass.: Harvard University Press
- Havránek, Bohuslav. 1964. 'Zum Problem der Norm in der heutigen Sprachwissenschaft und Sprachkultur'. In *A Prague school reader in linguistics*, 413–20. Indiana University studies in the history and theory of linguistics
- Heine, Bernd, and Tania Kuteva. 2005. *Language Contact and Grammatical Change*. Cambridge University Press
- . 2006. *The Changing Languages of Europe*. Oxford University Press
- . 2010. 'Contact and Grammaticalization'. In *The Handbook of Language Contact*, edited by Raymond Hickey, 86–105. Wiley-Blackwell
- Herrity, Peter. 1978. 'Puristic Attitudes in Serbia in the Second Half of the Nineteenth Century'. *The Slavonic and East European Review* 56 (2): 202–23
- Hickey, Raymond. 2010. *The Handbook of Language Contact*. Wiley-Blackwell
- Hopper, Paul J., and Elizabeth Closs Traugott. 2003. *Grammaticalization*. 2. ed. Cambridge Textbooks in Linguistics. Cambridge University Press
- Hornberger, Nancy H. 2006. 'Frameworks and Models in Language Policy and Planning'. In *An Introduction to Language Policy, Theory and Method*, edited by Thomas Ricento, 36–53. Blackwell Pub
- Horvat, Marijana, and Barbara Štebih-Golub. 2010. 'Posljedice internacionalizacije u hrvatskome jeziku'. *Rasprave Instituta za Hrvatski Jezik i Jezikoslovlje* 36 (1): 1–21

- Hudeček, Lana, and Milica Mihaljević. 2009. *Jezik medija - publicistički funkcionalni stil*. Zagreb: Hrvatska sveučilišna naklada.  
[http://www.superknjizara.hr/?page=knjiga&id\\_knjiga=29556](http://www.superknjizara.hr/?page=knjiga&id_knjiga=29556).
- Hudeček, Lana, Maja Matković, and Igor Ćutuk. 2011. *Jezični priručnik Coca-Cole HBC Hrvatska - Hrvatski jezik u poslovnoj komunikaciji*. 1. Zagreb: Coca-Cola HBC Hrvatska
- IHJJ, Institut za hrvatski jezik i jezikoslovlje. 2015a. *Jezični Savjeti - \*Internet-Kafić > Internetski Kafić*. Accessed September 23.  
<http://savjetnik.ihjj.hr/savjet.php?id=214>.
- . 2015b. *Jezični Savjeti - \*Multi- > Više-*. Accessed September 23.  
<http://savjetnik.ihjj.hr/savjet.php?id=273>.
- . 2015c. *Jezični Savjeti - \*Donor > donator > Darivatelj*. Accessed September 23. <http://savjetnik.ihjj.hr/savjet.php?id=125>.
- . 2015d. *Jezični Savjeti -Imenice Koje Završavaju Na -Itet*. Accessed September 23. <http://savjetnik.ihjj.hr/savjet.php?id=241>.
- . 2016. *O Struni | Struna | Hrvatsko Strukovno Nazivlje*. Accessed January 30. <http://struna.ihjj.hr/page/o-struni/#povijest>.
- Ivić, Milka. 1971. 'Leksema jedan i problem neodređenog člana'. *Zbornik Matice srpske za filologiju i lingvistiku* XIV (1): 103–20
- Ivić, Milka. 1965. 'Problem norme u književnom jeziku (Referat na Petom kongresu jugoslavenskih slavista u Sarajevu)'. *Jezik: časopis za kulturu hrvatskoga književnog jezika* 13 (1): 1–8
- Ivić, Pavle. 1990. *O jeziku nekadašnjem i sadašnjem*. Beograd : Priština: BIGZ / JEDINSTVO
- . 1992. 'Language Planning in Serbia Today'. In *Language Planning in Yugoslavia*, edited by Celia Hawkesworth and Ranko Bugarski, 101–10. Columbus, Ohio: Slavica Pub
- . 1998. *Pregled istorije srpskog jezika*. Pavle Ivić - Celokupna dela, VIII. Izdavačka knjižarnica Zorana Stojanovića
- Ivić, Pavle, Ivan Klajn, Mitar Pešikan, and Branislav Brborić. 2011. *Srpski jezički priručnik*. 5th edition. Beograd: Beogradska knjiga
- Jacobsen, Per. 1989. *Litteraturen i Jugoslavien, [Bind] 2 , Serbisk litteratur 1770-1945*. Materialer Københavns Universitets Institut for Slavistik og Øststatsforskning 7. CARitzel
- Jacobsen, Per. 2010. 'U potrazi za idealnom formom: o jezičkom purizmu u Evropi'. In *Južnoslovenske teme*, by Per Jacobsen, 111–26. Belgrade: SlovoSlavia
- Jakobson, Roman. 1959. 'On Linguistic Aspects of Translation'. In *On Translation*, edited by Reuben A. Brower, 23:232–39. Harvard Studies in Comparative Literature. Harvard: Harvard University Press
- Jonke, Ljudevit. 1965. 'Problem norme u hrvatskosrpskom književnom jeziku'. *Jezik: časopis za kulturu hrvatskoga književnog jezika* 13 (1): 8–14
- . 1971. *Hrvatski književni jezik 19. i 20. stoljeća*. Matica hrvatska
- Joseph, John Earl. 1987. *Eloquence and Power, the Rise of Language Standards and Standard Languages*. Open Linguistics Series. Pinter

- . 2006. *Language and Politics*. Edinburgh Textbooks in Applied Linguistics. Edinburgh University Press
- Jovanović, Vladan Z. 2013. 'O rečima sa vezanim osnovama stranog porekla u srpskom jeziku iz ugla gramatike (tvorbe reči) i leksikografije'. *Intralingvistički i ekstralingvistički činioci u formiranju i razvoju srpskog standardnog jezika*, NSSVD – MSC 42 (1): 295–308
- Jojić, Željko, Goranka Blagus Bartolec, Lana Hudeček, Kristian Lewis, Milica Mihaljević, Ermina Ramadanović, Matea Birtić, et al. 2013. *Hrvatski pravopis*. Zagreb: Institut za hrvatski jezik i jezikoslovlje
- Hraste, Mate, J. Vuković, and R. Aleksić. 1960. *Pravopis srpskohrvatskog književnog jezika*. Novi Sad, Zagreb: Matica srpska, Matica hrvatska
- Kaplan, Robert B., and Richard B. Baldauf. 1997. *Language Planning from Practice to Theory*. Multilingual Matters ; 108. Multilingual Matters
- Katičić, Radoslav. 1997. 'Undoing A "unified" language: Bosnian, Croatian, Serbian'. In *Undoing and Redoing Corpus Planning*, edited by Michael Clyne, 165–91. Berlin: Mouton de Gruyter
- . 2008. 'On "unified Serbo-Croatian" and Its "new Successor Languages": A Case Study of Scholarly Treatment'. *International Journal of the Sociology of Language* 2008 (191): 27–44
- Katunar, Daniela, Jana Willer Gold, and Tena Gnjatović. 2013. 'Achieving Specificity in an Articleless Language: Specificity Markers in Croatian'. *Suvremena Lingvistika* 39 (75): 25–43
- Kekez, Stipe. 2012. 'Naziv za mjesto gdje se prodaju pite'. Edited by Sanda Ham. *Jezik, časopis za kulturu hrvatskoga književnog jezika* 59 (2): 53–59
- Klaić, Bratoljub. 1982. *Rječnik stranih riječi*. Zagreb: Nakladni zavod Matice hrvatske
- Klajn, Ivan. 1996. 'Leksika'. In *Srpski jezik na kraju veka*, edited by Milorad Radovanović, 37–86. Institut za srpski jezik SANU: Službeni glasnik
- . 2002. *Tvorba reči u savremenom srpskom jeziku*. - *Prvi deo. Slaganje prefiksacija*. Vol. 1. 2 vols. Tvorba reči u savremenom srpskom jeziku. Beograd: Zavod za udžbenike i nastavna sredstva
- . 2003. *Tvorba reči u savremenom srpskom jeziku*. - *Drugi deo. Sufiksacija i konverzija*. Vol. 2. 2 vols. Beograd: Institut za srpski jezik SANU
- . 2008. 'Purizam i antipurizam u današnjem srpskom jeziku'. *Južnoslovenski Filolog* 64: 153–176.
- . 2011. 'Pravopisni problemi i nedoumice'. In *Srpski jezički priručnik*, 5th edition, 93–118. Beograd: Beogradska knjiga
- Kloss, Heinz. 1967. "'Abstand Languages" and "Ausbau Languages"'. *Anthropological Linguistics* 9 (7): 29–41
- . 1969. *Research Possibilities on Group Bilingualism: A Report*. Quebec: Centre international de recherches sur le bilinguisme.
- . 1976. 'Abstandsparchen und Aubauprachen'. Edited by Joachim Göschel, Norbert Nail, and Gaston van der Elst. *Zur Theorie des*

- Dialekts, Aufsätze aus 100 Jahren Forschung mit biographischen Anmerkungen zu den Autoren*, Zeitschrift für Dialektologie und Linguistik Beihefte. Neue Folge 16,
- Koneski, Blaže. 1980. 'Macedonian'. In *The Slavic Literary Languages, Formation and Development.*, edited by Alexander M. Schenker, Edward Stankiewicz, and Micaela S. Iovine, 53–63. Yale: Yale Concilium on International and Area Studies
- Kordić, Snježana. 2010. *Jezik i nacionalizam*. Zagreb: Durieux
- Košutar, Petra, and Branka Tafra. 2009. 'Rječotvorni modeli u hrvatskom jeziku'. *Suvremena lingvistika*, no. 67: 87–107
- Kristiansen, Tore. 2001. 'Standardsprog : om termens introduktion og brug i Danmark'. In *Sproglige åbninger, E som Erik, H som 70, festskrift til Erik Hansen 18. september 2001*, 347–57. Hans Reitzel
- Kulick, Don. 1992. *Language Shift and Cultural Reproduction, Socialization, Self and Syncretism in a Papua New Guinean Village*. Studies in the Social and Cultural Foundations of Language 14. Cambridge University Press
- Langston, Keith, and Anita Peti-Stantić. 2003. 'Attitudes towards Linguistic Purism in Croatia: Evaluating Efforts at Language Reform'. In *At War with Words*, edited by Mirjana N. Dedać and Daniel N. Nelson, 247–82. Language, Power and Social Process 10. Berlin, New York: De Gruyter
- . 2011. 'A Language Academy by Any Other Name(s): The Case of Croatia'. *Language Policy* 10 (4): 343–60
- . 2014. *Language Planning and National Identity in Croatia*. Palgrave Studies in Minority Languages and Communities. Palgrave Macmillan
- Laycock, D. C. 2001. 'Linguistic Diversity in Melanesia - A Tentative Explanation'. In *Ecolinguistics Reader : A Selection of Articles on Language, Ecology and Environment*, 167–71. London, GBR: Continuum International Publishing
- Leach, Edmund Ronald. 1964. *Political Systems of Highland Burma*. London, GBR: University of London, The Athlone Press
- Lehmann, Christian. 1985. 'Grammaticalization: Synchronic variation and diachronic change'. *Lingua e stile* 20: 303–18
- Lenček, Rado. 1976. 'A Few Remarks for the History of the Term "Serbo-croatian" language'. *Matica srpska. Zbornik za filologiju i lingvistiku* 19: 45–53
- Leto, Maria Rita. 2004. 'Danica ilirska i pitanje hrvatskoga književnog jezika'. *Slavica tergestina*, Studia slavica, 3 (11–12): 163–90
- Lieber, Rochelle. 2011. 'A Lexical Semantic Approach to Compounding'. In *The Oxford Handbook of Compounding*, edited by Rochelle Lieber and Pavol Štekauer, 78–104. Oxford Handbooks in Linguistics. Oxford University Press

- Lieber, Rochelle, and Pavol Štekauer, eds. 2011. *The Oxford Handbook of Compounding*. Oxford Handbooks in Linguistics. Oxford University Press
- Lindstrom, Nicole. 2003. 'Yugonostalgia: Restorative and Reflective Nostalgia in Former Yugoslavia'. *Slavic Review* 62 (1): 231–42
- Lipka, Leonhard. 1992. *An Outline of English Lexicology Lexical Structure, Word Semantics, and Word-Formation*. 2nd. edition. Tübingen: Max Niemeyer Verlag
- Liu, Xiaofeng, ed. 2015. *Sino-Theology and the Philosophy of History: A Collection of Essays by Liu Xiaofeng*. BRILL
- Lončarić, Mijo, ed. 1998. *Hrvatski jezik*. Opole, Poland: Uniwersytet Opolski, Instytut Filologii Polskiej
- Lyons, Chris. 1999. *Definiteness*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press
- Maretić, Tomislav. 1924. *Hrvatski ili srpski jezični savjetnik za sve one, koji žele dobro govoriti i pisati književnim našim jezikom* 'Rjecnik hrvatskoga jezika', *Dopuna Broz-Ivekovićevu 'Rječniku hrvatskoga jezika'*. Znanstvena djela za opću naobrazbu 7
- . 1963[1899]. *Grammatika hrvatskoga ili srpskoga književnog jezika*. 3rd edition. Zagreb: Matica hrvatska
- Maretić, Tomo. 1914. Section 10 in 'Anketa o južnom ili istočnom narečju'. *Srpski književni glasnik* XXXII (2): 222–24
- Matasović, Ranko. 2011. 'Što je srednjojužnoslavenski?' *Jezik* 58 (3): 81–90
- Mathesius, Vilém. 1932. 'O potřebě stability ve spisovném jazyce'. In *Spisovna čestina a jazyková kultura*, edited by Boh Havránek and Miloš Weingart, 14–31
- Matić, Gordana D. 2012. 'Прилог развоју модела стандардизације терминологије менаџмента у организацији'. PhD thesis, Beograd: University of Belgrade
- Matras, Yaron. 2009. *Language Contact*. Cambridge University Press
- . 2010. 'Contact, Convergence, and Typology'. In *The Handbook of Language Contact*, edited by Raymond Hickey, 66–85. Wiley-Blackwell
- Merriam-Webster Dictionary. 2015. March 3. <http://www.merriam-webster.com>.
- Mićanović, Krešimir. 2006. 'Mjesto standardologije u jezikoslovnoj kroatistici'. In *Hrvatski jezik u XX. stoljeću*, edited by Samardžija, Marko and Ivo Pranjkovic, 559–85
- . 2012. 'Osamanaesto stoljeće u hrvatskoj jezičnoj povijesti: standardološka perspektiva'. *Povijest hrvatskoga jezika - Književnost i kultura devedesetih*, Zbornik radova Zagrebačke slavističke škole, , 97–113
- Mihaljević, Milica. 2002. 'Internetsko nazivlje u publicističkome funkcionalnom stilu'. *Suvremena Lingvistika* 53–54: 226–28

- . 2012. 'Ekoškola, eko škola ili eko-škola?' *Jezik: Periodical for the Culture of the Standard Croatian Language* 59 (2): 71–72
- Mihaljević, Milica, and Barbara Štebih-Golub. 2009. 'Hrvatsko i srpsko računalno nazivlje'. In *Die Unterschiede zwischen dem Bosnischen/Bosniakischen, Kroatischen und Serbische - Lexik - Wortbildung - Phraseologie*, edited by Branko Tošović, 2:137–42. Graz: LIT Verlag Münster
- Milroy, James. 2001. 'Language Ideologies and the Consequences of Standardization'. *Journal of Sociolinguistics* 5 (4): 530–55..
- Moguš, Milan. 1995. *A History of the Croatian Language: Toward a Common Standard*. Zagreb: Globus
- Mønnesland, Svein. 1999. *Før Jugoslavia og etter*. 4th ed. Oslo: Syress
- . 2002. *Bosnisk, kroatisk, serbisk grammatikk*. Syress Forlag
- Moser, Michael. 2013. 'The Fall of the Iron Curtain and the New Linguistic Landscape of East-Central Europe'. In *The Fall of the Iron Curtain and the Culture of Europe*, edited by Peter I. Batra, 104–17. Routledge
- Mrazović, Pavica. 2009. *Gramatika srpskog jezika za strance*. 2. Sremski Karlovci; Novi Sad: Izdavačka knjižarnica Zorana Stojanovića
- Mrazović, Pavica, and Zora Vukadinović. 1990. *Gramatika srpskohrvatskog jezika za strance*. Sremski Karlovci; Novi Sad: Izdavačka knjižarnica Zorana Stojanovića; Dobra vest
- Muhvić-Dimanovski, Vesna. 1992. 'Prevedenice - Jedan oblik neologizama'. *Rad Hrvatske akademije znanosti i umjetnosti* 446: 93–205
- . 2005. *Neologizmi - problemi teorije i primjene*. Zagreb: FF Press, Zavod za lingvistiku Filozofskog fakulteta Sveučilišta u Zagrebu
- Murray, Robert W. 1997. 'Historical Linguistics: The Study of Language Change'. In *Contemporary Linguistics, an Introduction*, edited by William D. O'Grady, Michael Dobrovolsky, and Mark Aronoff, 3., U.S. prepared by Mark Aronoff, 313–71. StMartin's Press
- Naylor, Kenneth. 1980. 'Serbo-Croatian'. In *The Slavic Literary Languages, Formation and Development*, edited by Alexander M. Schenker, Edward Stankiewicz, and Micaela S. Iovine, 65–83. Yale: Yale Concilium on International and Area Studies
- Nedergaard Thomsen, Ole. 2006. 'Towards an Integrated Functional-Pragmatic Theory of Language and Language Change'. In *Competing Models of Linguistic Change, Evolution and beyond*, edited by Ole Nedergaard Thomsen, 307–37. Amsterdam/NL: John Benjamins Publishing Company
- Neustupný, J.V. 1989. 'Language Purism as a Type of Language Correction'. In *The Politics of Language Purism*, edited by Björn H. Jernudd and Michael J. Shapiro, 211–24. Berlin, Boston: De Gruyter Mouton.
- 'Nobel Prize in Literature 1961'. 2016. Accessed July 7. [http://www.nobelprize.org/nobel\\_prizes/literature/laureates/1961/](http://www.nobelprize.org/nobel_prizes/literature/laureates/1961/).



- Nuorluoto, Juhani. 2012. *De slaviska standardspråkens framväxt : idéhistoriska och språksociologiska kontexter och koncept*. Uppsala, Sweden: Institutionen för moderna språk.
- Nørgård-Sørensen, Jens, Lars Heltoft, and Lene Schøsler. 2011. *Connecting Grammaticalisation*. Studies in Functional and Structural Linguistics 65. John Benjamins Publishing Company
- Oczkova, Barbara. 2010. *Hrvati i njihov jezik*. Zagreb: Školska knjiga
- Opačić, Nives. 2002. 'Strah od materinskoga (hrvatskoga) standardnog jezika'. *Suvremena Lingvistika* 2002 (53–54): 229–31
- . 2007. 'Prodor engleskih riječi u hrvatski jezik'. Edited by Sanda Ham. *Jezik, časopis za kulturu hrvatskoga književnog jezika* 54 (1): 22–27
- Otašević, Đorđe. 1998. 'Nova hibridna leksika u srpskom i ruskom jeziku'. *Slavistika* 2 (2): 99–102
- Pervaz, Draginja. 1996. 'Engleske skraćenice u srpskom jeziku'. In *O leksičkim pozajmljenicama*, edited by Judita Plankoš, 105–19. Subotica - Beograd: Institut za srpski jezik SANU, Gradska biblioteka Subotica
- Pešikan, Mitar, Jovan Jerković, and Mato Pižurica. 1993. *Pravopis srpskoga jezika*. Novi Sad: Matica srpska
- . 2010. *Pravopis srpskoga jezika*. Novi Sad: Matica srpska
- Peti-Stantić, Anita. 2008. *Jezik naš i/ili njihov : vježbe iz poredbene povijesti južnoslavenskih standardizacijskih procesa /*. Zagreb : Srednja Europa,
- Petrov, Bisser. 2011. 'Ideology and Education in Occupied Serbia 1941–1944'. *Etudes Balkaniques*, no. 1: 88–108
- Petrović, Dragoljub. 1996. 'Fonetika'. In *Srpski jezik na kraju veka*, edited by Milorad Radovanović, 87–110. Institut za srpski jezik SANU: Službeni glasnik
- Piper, Predrag, and Ivan Klajn. 2013. *Normativna gramatika srpskog jezika*. Novi Sad, Serbia: Matica srpska
- Plankoš, Judita, ed. 1996. *O leksičkim pozajmljenicama: zbornik radova sa naučnog skupa Strane reči i izrazi u srpskom jeziku sa osvrtom na isti problem u jezicima nacionalnih manjina (Gradska biblioteka, Subotica, 18-20. oktobar 1995)*. Subotica - Beograd: Institut za srpski jezik SANU, Gradska biblioteka Subotica
- Popović, Ljudmila. 2009. 'Leksičke inovacije u elektronskom diskursu srpskog i hrvatskog jezika'. In *Die Unterschiede zwischen dem Bosnischen/Bosniakischen, Kroatischen und Serbische - Lexik - Wortbildung - Phraseologie*, edited by Branko Tošović, 2:183–203. Graz: LIT Verlag Münster
- Popović, Mihailo. 1997. 'O nekim novim slučajevima polisemijskog semantičkog pozajmljivanja'. *Srpski jezik* 1/2 (2): 347–54
- Pranjeković, Ivo. 1997. *Jezikoslovna sporenja*. Zagreb: Konzor

- . 2000. 'Normative und paranormative Neuerungen in der kroatischen Sprache'. Edited by Barbara Kunzmann-Müller. *Die Sprachen Südosteuropas heute, Umbrüche und Aufbruch*, 66–74
- . 2006. 'Hrvatski jezik od godine 1945. do 2000.' In *Hrvatski jezik u XX. stoljeću*, edited by Marko Samardžija and Ivo Pranjković. Zagreb: Matica hrvatska
- . 2008. *Sučeljavanja*. Zagreb: Disput
- Prčić, Ljubica. 1996. 'Karakterizatori stranog porekla u srpskom jeziku'. In *O leksičkim pozajmljenicama*, edited by Judita Plankoš, 147–52. Subotica - Beograd: Institut za srpski jezik SANU, Gradska biblioteka Subotica
- Prčić, Tvrtko. 2006. 'Srpski ili anglosrpski - pitanje je sad'. In *Srpski jezik u normativnom ogledalu : 50 odluka Odbora za standardizaciju srpskog jezika /*, edited by Branko Brborić et al, 412–22. Biblioteka Put u reci. Beograd : Beogradska knjiga,
- Projekat Rastko. 2015. 'Odbor za standardizaciju srpskog jezika'. *Projekat Rastko - Biblioteka srpske kulture na internetu*. September 2. [http://www.rastko.rs/filologija/odbor/sporazum-potpisnici\\_c.html](http://www.rastko.rs/filologija/odbor/sporazum-potpisnici_c.html).
- Radovanović, Milorad. 1996. *Srpski jezik na kraju veka*. Institut za srpski jezik SANU: Službeni glasnik
- . 2000. 'From Serbo-Croatian to Serbian'. *Multilingua - Journal of Cross-Cultural and Interlanguage Communication* 19 (1–2): 21–36
- . 2003. *Sociolingvistika*. Biblioteka THEORIA. Sremski Karlovci; Novi Sad: Izdavačka knjižarnica Zorana Stojanovića
- . 2004. *Planiranje jezika i drugi spisi*. Biblioteka ELEMENTI 56. Sremski Karlovci; Novi Sad: Izdavačka knjižarnica Zorana Stojanovića
- . 2009. 'Srpski jezički standard'. In *Jezični varijeteti i nacionalni identiteti, prilozi proučavanju standardnih jezika utemeljenih na štokavštini*, edited by Lada Badurina, Ivo Pranjković, and Josip Silić, 199–213. Zagreb: Disput
- Romaine, Suzanne. 2004. 'Language-Contact Studies'. In *Sociolinguistics. An International Handbook of the Science of Language and Society*, 2nded., 1:49–58. Handbooks of Linguistics and Communication Science 3. Berlin/New York: De Gruyter
- Rožić, Vatroslav. 1904. *Barbarizmi u hrvatskom ili srpskom jeziku*. Zemun: Tiskara Jovana Puljo
- Samardžija, Marko. 1998. 'Leksik'. In *Hrvatski jezik*, edited by Mijo Lončarić, 133–52. Opole: Uniwersytet Opolski, Instytut Filologii Polskiej
- . 2008. *Hrvatski jezik, pravopis i jezična politika u Nezavisnoj Državi Hrvatskoj*. Zagreb: Hrvatska sveučilišna naklada
- . 2012. *Hrvatski jezik i pravopis od ujedinjenja do kraja Banovine Hrvatske (1918. - 1941.)*. Zagreb: Školska knjiga

- Sandøy, Helge, and Bente Selback. 2007. 'Avispråk og importord'. In *Fire dagar i nordiske aviser, ei jamføring av påverknaden i ordforrådet i sju språksamfunn*, edited by Bente Selback, 9–24. Novus
- 'Sava Centar'. 2016. *Sava Centar*. Accessed October 10. <http://www.sava-centar.net/index.php/en/sc/o-nama>.
- Scalise, Sergio. 1986. *Generative Morphology*. 2. ed. 18. Foris Publications
- Scalise, Sergio, and Antonietta Bisetto. 2011. 'The Classification of Compounds'. In *The Oxford Handbook of Compounding*, edited by Rochelle Lieber and Pavol Štekauer, 34–53. Oxford University Press
- Scholte, Jan Aart. 1999. *Globalization, a Critical Introduction*. London, Basingstoke: Macmillan, Macmillan Press.
- Silić, Josip. 2001. 'Hrvatski i srpski jezik'. In *Zbornik Zagrebačke slavističke škole: trideset godina rada : 1972-2001*, edited by Stipe Botica, 147–55. Zagreb: Filozofski fakultet, Zagrebačka slavistika škola
- . 2006. *Funkcionalni stilovi hrvatskoga jezika*. Zagreb: Disput
- Skelin Horvat, Anita. 2004. 'Posuđivanje u hrvatski jezik u dvama razdobljima'. *Suvremena lingvistika*, no. 57–58: 93–104
- Skok, Petar. 1955. 'O sufiksima -isati, -irati i -ovati'. *Jezik: Periodical for the Culture of the Standard Croatian Language* 4 (2): 36–43
- Slijepčević, Svetlana. 2013. 'O složenim jedinicama sa stranim indeklinabilnim prepozitivnim determinatorom'. *Intralingvistički i ekstralingvistički činioci u formiranju i razvoju srpskog standardnog jezika*, NSSVD – MSC 42 (1): 321–30
- 'Službeni glasnik RS. 2010. 'Zakon o službenoj upotrebi jezika i pisama'. Official Gazette of RS, No. 30/2010
- Sočanac, Milica. 2009. 'Anglicizmi u hrvatskom i srpskom jeziku: sličnosti i razlike'. In *Die Unterschiede zwischen dem Bosnischen/Bosniakischen, Kroatischen und Serbische - Lexik - Wortbildung - Phraseologie*, edited by Branko Tošović, 2:239–52. Graz: LIT Verlag Münster
- Sofronić, Zlatomir. 1974. *Ustav S(ocijalističke) F(ederativne) R(epublike) J(ugoslavije)*. Beograd: Prosveta
- Southerland, Ronald H., and Francis Katamba. 1997. 'Language in Social Contexts'. In *Contemporary Linguistics, an Introduction*, edited by William D. O'Grady, Michael Dobrovolsky, and Mark Aronoff, 3., U.S. prepared by Mark Aronoff, 540–90. St Martin's Press
- 'Sprachliche Varietäten'. 2015. *Eugenio Coseriu – Archiv*, Eberhard Karls Universität Tübingen, February 13. <http://www.romling.uni-tuebingen.de/coseriu/variet.html>.
- Srpski jezički atelje. 2014. 'Pravopis srpkoga jezika'. *Srpski jezički atelje*. <http://www.srpskijezickiatelje.com/knjige:1>.
- Stanković, Bogoljub. 1997. 'Sintagmatska inoviranja u jeziku štampe'. *Južnoslovenski filolog* LIII: 87–97
- Starčević, Anđel. 2006. 'Imenice kao atributi - nove strukture u hrvatskom jeziku'. In *Jezik i mediji - jedan jezik: više svjetova : zbornik*, edited

- by Jagoda Granić, 645–55. Zagreb-Split: Hrvatsko društvo primijenjene lingvistike
- Stevanović, Mihailo. 1964. *Savremeni srpskohrvatski jezik (Gramatički sistemi i književnojezička norma)*. Vol. 1. 2 vols. Beograd: Naučno delo
- Stewart, William A. 1968. 'A Sociolinguistic Typology for Describing National Multilingualism'. The Hague: Mouton
- Stijović, Rada. 2009. 'O leksici srpskog književnog jezika posle raskida srpskohrvatske jezičke zajednice'. In *Die Unterschiede zwischen dem Bosnischen/Bosniakischen, Kroatischen und Serbische - Lexik - Wortbildung - Phraseologie*, edited by Branko Tošović, 2:215–21. Graz: LIT Verlag Münster
- Sundhaussen, Holm. 1973. 'Der Einfluss der Herderschen Ideen auf die Nationsbildung bei den Völkern der Habsburger Monarchie'. Disputats, München
- . 2009. *Istorija Srbije : od 19. do 21. veka*. Beograd: Clio
- Šević, Radmila B. 1996. 'Srpski i engleski - jezici u kontaktu?' In *O leksičkim pozajmljenicama*, edited by Judita Plankoš, 81–88. Subotica - Beograd: Institut za srpski jezik SANU, Gradska biblioteka Subotica
- Šipka, Danko. 2010. 'Jezik srpskih i hrvatskih medija - leksičko-semantički aspekt'. In *Die Unterschiede zwischen dem Bosnischen/Bosniakischen, Kroatischen und Serbische - Grammatik*, edited by Branko Tošović, 3:459–68. Graz: LIT Verlag Münster
- Šipka, Milan. 2005. 'Standardni jezik i jezička politika u Bosni i Hercegovini 1918.-1970.' In *Jezik u Bosni i Hercegovini*, edited by Svein Mønnesland, 407–34. Oslo
- . 2006. 'Za kodeks upotrebe srpskog standardnog jezika'. In *Srpski jezik u normativnom ogledalu : 50 odluka Odbora za standardizaciju srpskog jezika* / edited by Branko Brborić et al, 294–313. Biblioteka Put u reci. Beograd : Beogradska knjiga
- Škiljan, Dubravko. 1988. *Jezična politika*. Naprijed
- . 2000. 'From Croato-Serbian to Croatian: Croatian Linguistic Identity'. *Multilingua - Journal of Cross-Cultural and Interlanguage Communication* 19 (1–2): 3–20.
- Škodrić, Ljubinka. 2011. 'Education as Part of General Nedić's Government the Ideology'. *Istorija 20. Veka* 29 (1): 151–62
- Šonje, Jure, ed. 2000. *Rječnik hrvatskoga jezika*. Zagreb: Školska knjiga.
- Štebih, Barbara. 2008. 'Morfološka adaptacija posuđenica'. *Suvremena Lingvistika* 66 (2): 243–59
- Štekauer, Pavol, and Rochelle Lieber. 2005. *Handbook of Word-Formation*. Vol. 64. Studies in Natural Language and Linguistic Theory. Dordrecht, The Netherlands: Springer
- Štekauer, Pavol, Salvador Valera, and Livia Körtyvélyessy. 2012. *Word-Formation in the World's Languages a Typological Survey*. Cambridge: Cambridge University Press

- Tadić, Marko. 2009. 'New Version of the Croatian National Corpus'. Edited by Dana Hlaváčková, Aleš Horák, Klara Osobě, and Pavel Rychlý. *After Half a Century of Slavonic Natural Language Processing*, 219–28
- 'Nobel Prize in Literature 1961'. 2016. Accessed July 7. [http://www.nobelprize.org/nobel\\_prizes/literature/laureates/1961/](http://www.nobelprize.org/nobel_prizes/literature/laureates/1961/).
- Thomas, George. 1991. *Linguistic Purism*. Studies in Language and Linguistics. Longman
- . 1992. 'Lexical Purism as an Aspect of Language Cultivation in Yugoslavia'. In *Language Planning in Yugoslavia*, edited by Ranko Bugarski and Celia Hawkesworth. Columbus, Ohio: Slavica Pub
- Thomas, Paul-Louis. 2002. 'Le serbo-croate (bosniaque, croate, monténegrin, serbe) : de l'étude d'une langue à l'identité des langues'. *Revue des études slaves* 74 (2): 311–25
- Thomason, Sarah. 2001. *Language Contact, an Introduction*. Edinburgh University Press
- . 2003. 'Contact as a Source of Language Change'. In *The Handbook of Historical Linguistics*, edited by Brian D. Joseph and Richard D. Janda, 687–712. Blackwell Publishing Ltd
- Tošović, Branko. 2002. *Funkcionalni stilovi (Funktionale Stile)*. Graz: Institut für Slawistik der Karl-Franzens-Universität, Graz
- . , ed. 2009. *Die Unterschiede Zwischen Dem Bosnischen/Bosniakischen, Kroatischen Und Serbische - Lexik - Wortbildung - Phraseologie*. Vol. 2. 3 vols. Graz: LIT Verlag Münster
- . 2010. 'Die grammatikalischen Unterschiede zwischen dem Bosnischen/Bosniakischen, Kroatischen und Serbischen'. In *Die Unterschiede zwischen dem Bosnischen/Bosniakischen, Kroatischen und Serbische - Grammatik*, edited by Branko Tošović, 3:131–87. Graz: LIT Verlag Münster
- Turk, Marija, and Maja Opašić. 2008. 'Linguistic Borrowing and Purism in the Croatian Language'. *Suvremena Lingvistika* 33 (65): 73–88
- Utvić, Miloš. 2011. 'Annotating The Corpus of Contemporary Serbian'. *IN-FOtheca* XII (2): 36a–47a
- . 2014. 'Liste učestanosti korpusa savremenog srpskog jezika'. *Srpski jezik i njegovi resursi: Teorija, opis i primene* 43 (3): 241–62
- Vasić, Vera, ed. 1990. *Jezička politika i planiranje jezika u Jugoslaviji*. Novi Sad: Institut za južnoslovenske jezike, Filozofski fakultet u Novom Sadu
- Vasyl'eva, Milica. 2009. 'Dinamika razvoja srpskog, hrvatskog i bosanskog leksika krajem XX. - početkom XXI. stoljeća'. In *Die Unterschiede zwischen dem Bosnischen/Bosniakischen, Kroatischen und Serbische - Lexik - Wortbildung - Phraseologie*, edited by Branko Tošović, 2:165–82. Graz: LIT Verlag Münster
- Veskić, J. 1932. *Rečnik slovenačko-srpsko-hrvatski*. Beograd: Štamparija Drag. Gregorić

- VNHSJ. 2013. 'Vijeće za normu hrvatskoga standardnog jezika'. *Jezik: časopis za kulturu hrvatskoga književnog jezika* 60 (2–4): 41–160
- Wikinfo, na hrvatskom jeziku. 2016. 'Dialektni imenski član'. Accessed January 28. [http://wwtrade.org/Hrvatski/index.php/Dialektni\\_imenski\\_%C4%8Dlan](http://wwtrade.org/Hrvatski/index.php/Dialektni_imenski_%C4%8Dlan).
- Winford, Donald. 2010. 'Contact and Borrowing'. In *The Handbook of Language Contact*, edited by Raymond Hickey, 170–87. Wiley-Blackwell
- Wonisch, Arno. 2009. 'Leksičke razlike u jeziku sporta u medijima BKS-područja'. In *Die Unterschiede zwischen dem Bosnischen/Bosniakischen, Kroatischen und Serbische - Lexik - Wortbildung - Phraseologie*, edited by Branko Tošović, 2:205–13. Graz: LIT Verlag Münster
- Wright, Sue. 2004. *Language Policy and Language Planning, from Nationalism to Globalisation*. Palgrave Macmillan
- Zett, Robert. 1969. 'O problematici složenica tipa "nogomet"'. *Jezik, časopis za kulturu hrvatskosrpskoga književnog jezika* 16/1968/69 (4): 103–10
- . 1970. *Beiträge zur Geschichte der Nominalkomposita im Serbokroatischen, die altserbische Periode*. Slavistische Forschungen 9. Böhlau

## Resumé

Afhandlingen undersøger hvordan usus i henholdsvis kroatisk og serbisk ændres i både leksikalsk og grammatikalsk henseende, når sprogene og deres brugere udsættes for påvirkninger fra såvel andre sprog som sprogenes vogtere.

På basis af en introduktion til serbisk og kroatisk sprogs turbulente fælles historie, hovedsageligt under navnet serbokroatisk, samt til de sproglige træk der i dag anvendes til at understrege rigtigheden i at betragte serbisk og kroatisk som to adskilte sprog (kap. 1) undersøges det ved hjælp af sociolingvistiske teorier om sprogplanlægning (kap. 2) samt teorier om sprogforandring fra den historiske lingvistik og dens gren kontaktlingvistikken (kap. 3) hvorvidt påvirkningerne fra især engelsk, men også andre fremmedsprog, bidrager til at øge eller mindske den eksisterende forskel mellem hvad der tidligere officielt ansås som to varianter af serbokroatisk. Undersøgelsen (kap. 4) bygger, under inddragelse af eksisterende forskning, på empiriske undersøgelser af fem forskellige fænomener, der repræsenterer fire typer af kontaktmotiverede sprogforandringer:

- 1) Indlån af leksikalsk materiale (substantiver og substantivlignende bundne morfemer)
- 2) Indlån af grammatisk materiale (afledningsaffikser)
- 3) Indlån af orddannelsesmønstre (i sammensatte substantiver)
- 4) Indlån af bestemthedskategorien (i form af en grammatikalisering af talordet *jedan* (en/et) og det ubestemte pronomener *neki* (en eller anden), der gør at disse ord fungerer som ubestemte artikler).

Afslutningsvis (kap. 5) konkluderes det på den ene side at:

a) Iøjnefaldende påvirkninger udefra i form af frie leksikalske morfemer såsom substantiver, hjulpet på vej af sproglig purisme og antipurisme blandt såvel sprogbrugere som de sprognormerende aktører, bidrager til at kroatisk og serbisk sprogbrug differentieres.

På den anden side konkluderes det at:

b) Indlån af bundne morfemer, orddannelsesmønstre samt bestemthedskategorien har den modsatte effekt.

Denne åbenbare modsætning anses at bero på at serbisk og kroatisk deler et fælles underliggende sprogsystem, der er skjult for det blotte øje, hvorimod leksikalske indlån har en symbolsk værdi, der påvirker både den foreskrevne sprognorm og ususnormen.

## Summary

This PhD thesis is a comparative study of contact-induced linguistic innovations in Serbian and Croatian with a view to determine whether these innovations enhance or diminish the differences between Serbian and Croatian language usage. The empirical focus is on written language use from the last 20-30 years, during which Serbian and Croatian have existed as two distinct standard languages as opposed to their earlier status as one: Serbo-Croatian.

The analyses of empirical data (in Chapter 4) are informed by:

- Theoretical frameworks describing language planning (Chapter 2),
- General theories of language change and, more specifically, language change induced by contact with other languages (Chapter 3), and on:
- Recent research and general descriptions of the phenomena under investigation (Chapter 4).

In the introduction (Chapter 1), the modern history of the status of Serbian, Croatian and Serbo-Croatian put the strained relationship between Serbian and Croatian and their speakers into context.

I investigated five specific phenomena and four general types of contact-induced linguistic innovations:

- 1) Replication of lexical matter (nouns and semiwords)
- 2) Replication of grammatical matter (derivational affixes)
- 3) Replication of word-formational patterns (in compounds)
- 4) Replication of a grammatical pattern (viz. indefinite marking of NPs which includes a grammaticalisation of the numeral *jedan* ('one') and indefinite pronoun *neki* ('some'))

Finally, in chapter 5, I summarise my findings and conclude that:

- a) As far as the patterns, which are part of underlying structure of the languages, are concerned, the usage norms of Serbian and Croatian undergo identical contact-induced changes.
- b) As far as the grammatical matter, i. e. derivational affixes, is concerned the usage in Serbian and Croatian is more or less identical in spite of differing prescribed norms.
- c) On the surface of the languages, and when the innovation consists of replication of very visible free morphemes (nouns), the distance between Serbian and Croatian is increasing.