

1. General

1.1. State of the Art and Methodology

Numerous contemporary monolingual Bosnian / Croatian / Serbian (see §1.3) grammars are available, most importantly Babić et al. (1991), Barić et al. (1995), Mrazović and Vukadinović (1988), Stanojčić and Popović (1992), and Stevanović (1981). Recently several manuals of orthography appeared (which also treat some grammatical issues), such as: Babić et al. (1995), Nikčević (1993), Halilović (1996), Pešikan et al. (1994), Simić et al. (1993). Finally, monographs on various grammatical levels and narrower problems are available in Bosnian / Croatian / Serbian and other languages.

The work on Bosnian / Croatian / Serbian grammar in English falls into two major categories. There are various textbooks, such as Hawkesworth (1998), Norris (1995), Magner (1995), Šljivić-Šimšić (1985 and others), with limited grammar sketches. On the other hand, there are conference proceedings (such as those from the Formal Approaches to Slavic Languages annual conferences, see FDSL, 1997-, FASL, 1994-), where particular, oftentimes rather narrow, problems are addressed.

One should furthermore mention the outlines of the BCS language by Alexander (2006), Browne (1992), and Kordić (1997), as well as numerous BCS - English contrastive monographs from the Zagreb BCS - English Contrastive project headed by Filipović (1975-).

While all aforementioned sources constitute significant contributions to presenting BCS grammar to English speakers, the English learner of BCS would often have to resort to several different sources to obtain full information about grammatical phenomena. In addition, the learner is likely to be confronted with the following inadequacies of some resources: a) non-comprehensiveness, b) imprecise information, c) data distorted by one's linguistic approach or political attitude, d) information overload. That is to say, only major points of grammar are discussed in some sources without addressing any details, others are interspersed with imprecise statements such as "some nouns have a feature X", "most verbs follow the pattern Y." Some studies lack lists of exceptions, yet other resources omit data inconsistent with their theoretical or political viewpoint. Finally, some sources feature impenetrably intricate descriptions where the description per se is more prominent than the data described.

A manifest need therefore exists for a concise yet comprehensive BCS grammar intended primarily for beginning to intermediate English learners, which would provide a hands-on account of the BCS grammatical system as comprehensively as possible.

This grammar is laid out as a comprehensive yet user-friendly reference for beginning to intermediate English learners. The basic grammar text gives rules in a form of DECISION-TREES, tables and figures. Longer lists of

Chapter 1: General

exceptions are provided in the appendices. Most sections are divided into a STRUCTURAL DESCRIPTION and a CONTRASTIVE SECTION. Features that contrast with English are elaborated in detail. Along with the sections containing grammatical information in a narrower sense (PHONOLOGY with PROSODY, INFLECTIONAL and LEXICAL MORPHOLOGY, SYNTAX), the grammar contains practical METAGRAMMATICAL information such as that on the ORTHOGRAPHY and the use of PRAGMATIC OPERATORS, which the learner will need in everyday communication. Main differences between standard and substandard grammatical forms are also provided.

Three principal theoretical approaches deployed in the grammar include MINIMAL INFORMATION GRAMMAR, the application of DECISION THEORY in modeling language structures, and COGNITIVE LINGUISTICS. They were used in the first stage of preparing the text, followed by the application of a user friendly interface. Several experiments involving Slavic grammar structures have been conducted and described in Šipka 1999 and 2001. An overarching goal of this grammar is to equip its users with the heuristics which will enable a swift learning process and easy solutions to any problem in the use of the language.

1.1.1 Graphic and other conventions of the layout

The following conventions are used in presenting the material.

- In the main text, all Bosnian / Croatian / Serbian examples are provided in *italics* followed by their English equivalents in single quotes. Examples in the tables are given as plain text.
- The first appearance of all technical linguistic terms is followed by a non-technical explanation, usually in parentheses.
- Important technical terms are marked in SMALL CAPS and are defined in the grammatical index.
- All tables are numbered for easier reference.
- In all tables, the category names in the first row and column are italicized.
- All lists inside the text are either bulleted or lettered.
- Portions of the text considered to be a preserve of advanced users (e.g., trained linguists, language instructors, etc.) are marked using a vertical line on the left margin.
- Portions of the text which pertain to only one ethnic or geographic form of the language are marked with a scissor icon preceding and following the text.
- The asterisk is used in the phonology section to mark a wildcard, i.e., any single sound/character or a cluster thereof while its function in the syntax section is to mark unacceptable (UNGRAMMATICAL) sentences.

A Comparative Reference Grammar of Bosnian / Croatian / Serbian

1.2. Affiliation and Typological Characteristics

Bosnian / Croatian / Serbian is an INDO-EUROPEAN and SLAVIC LANGUAGE. Its GENETIC AFFILIATION (i.e., the historical links with other languages which are believed to have developed from the same source) is provided in Table 1.2.1.

<i>Family</i>	<i>Subfamily</i>	<i>Group</i>	<i>Subgroup</i>	<i>Language</i>	
Indo-European	Slavic	Southern	South-Western	Bosniac/Croatian/ Serbian	
				Slovenian	
				South-Eastern	Bulgarian
			Macedonian		
			Old Church Slavonic		
			Eastern	Russian	
		Western	Lechitic	Polish	
				Sorbian	Upper Sorbian
				Lower Sorbian	
			Czecho-Slovak	Czech	
				Slovak	
				[...]	
				[...]	
Germanic	Western	Northern	English		
			[...]		

Table 1.2.1 Genetic Affiliation of Bosnian / Croatian / Serbian

The similarity of Bosnian / Croatian / Serbian to the other languages of the Slavic group is substantial. These languages share a sizable amount of both lexical and grammatical features. Learning Bosnian / Croatian / Serbian thus enhances the learning of other Slavic languages and vice versa.

Although ultimately related with the English language through the Indo-European family, lexemes and common features are hardly recognizable by a learner and thus in general do not facilitate learning.

TYPOLOGICALLY (i.e., according to its structural features), Bosnian / Croatian / Serbian is an INFLECTIONAL predominantly SUFFIXAL LANGUAGE with a flexible SVO (subject-verb-object) WORD ORDER. This means that words have different forms (i.e., INFLECTIONS) marked mostly with suffixes (endings) and that typical SENTENCE pattern follows the subject-verb-object order. Bosnian / Croatian / Serbian inflectional patterns are by far much more

Chapter 1: General

complex than their English counterparts. In addition, unlike English, Bosnian / Croatian / Serbian is a PRO-DROP LANGUAGE, which means that it omits personal pronouns when a finite verb form is used. Thus, the default BCS translation of *I work* is just *Radim*. ('work-1stp.Sg.Pres.Tense) rather than *Ja radim*. 'I work', which is an emphasized form.

1.3. Language and Its Varieties

There are at least three different STANDARD FORMS (i.e., official forms of the language as they are used in news, schools, courts, etc.) of Bosnian / Croatian / Serbian (also known as SERBO-CROATIAN), distinguished ethnically or geographically: 1. SERBIAN, used by Serbs (8 million), 2. CROATIAN, spoken by Croats (4.5 million), 3. BOSNIAC (also known as BOSNIAN), used by Bosniacs (also known as Bosnian Muslims, 2 million). The vast majority of Montenegrins (600,000) use the Serbian standard, but recently there have been attempts to create Montenegrin, as a fourth standard within Bosnian / Croatian / Serbian. These standards are in official use in the countries and territories where their respective ethnic groups constitute the majority. Serbian is thus the official standard of Serbia and Montenegro and of the Republic of Srpska (a part of Bosnia-Herzegovina). Croatian is the official standard of Croatia and the regions of the Federation of Bosnia-Herzegovina (a part of Bosnia-Herzegovina) with a Croatian majority. And finally, Bosniac is the official standard of the regions of the Federation of Bosnia-Herzegovina with a Bosnian Muslim majority. These standards share grammatical structures (with some minor exceptions) while exhibiting lexical differences. Because of these national standards, some linguists deny the existence of Bosnian / Croatian / Serbian per se, claiming that each of these standards represents a distinct Slavic language (see Okuka, 1998, Bugarski and Hawkesworth, 1992 for more information). The term BOSNIAC vs. BOSNIAN is contested politically. In the Bosniac standard the term is *bosanski* (= BOSNIAN), in Croatian and Serbian standards, the term is *bošnjački* (= BOSNIAC).

The Serbian standard allows two pronounciational VARIANTS in some words: the EKAVIAN (which has an *e* in the stem of these words) and IJEKAVIAN (which has *ije,je*, or *i* in the stem of all these words instead of the *e* in Ekavian). For example, the Ekavian forms are *mleko* 'milk', *hteti* 'want-Infinitive', *hteo* 'want-Past Participle') and their Ijekavian counterparts *mlijeko*, *htjeti*, *htio*. The Ekavian and Ijekavian pronunciation are not interchangeable; a text must be either Ekavian or Ijekavian. The Croatian and Bosniac standards use solely the Ijekavian pronunciation.

The division within the COLLOQUIAL field is an even more intricate one. In addition to being divided into colloquial in its narrower sense (i.e., general colloquial sphere, the way speakers use the language in unofficial communication), DIALECTS (i.e., rural vernaculars, the manner of speaking in the countryside), and SLANG (i.e., urban vernaculars of distinct social groups, e.g., drug users, criminals, school students, etc.), as well as according to the ethnic background, other factors also contribute to diversity in this field.

The factor of ethnicity is often interwoven with that of territory. Some lexemes are thus used exclusively by Serbs (such as *buva* 'lie, false information, joke'), while others are used only by speakers of Bosnian / Croatian / Serbian in Serbia (for example *bioskop* 'cinema'). Others are used by inhabitants of Bosnia (like *bolan* which is an intensifier used to address a single male in direct informal communication), notwithstanding their ethnic background. These territorial boundaries are not always consistent with either dialectal or administrative borders, but more frequently reflect historic and/or cultural circles and influences.

The colloquial LEXICON is also divided along the lines of different social and professional groups, such as drug addicts, criminals, students, etc. Each of these groups constitutes a subculture of its own, with its own specific beliefs and values.

There is no one-to-one correspondence between the three ethnic standards and colloquial forms used by Serbs, Croats and Bosniacs. On the one hand, two urban vernaculars within the same ethnic group might be mutually unintelligible (as Croatian vernaculars of Split and Zagreb can be), and on the other hand, one urban vernacular can be shared by people with different ethnic affiliations (as was the case in pre-1990 Sarajevo). Thus, according to Playboy (1999) a policeman is *drot, pajkan, rijamu, rora* in Zagreb and *murjak, pandur, tovarić* in Split. In contrast, all ethnic groups in Sarajevo use the word *drot* to refer to a police officer. Furthermore, even if a colloquial lexeme is ethnically marked, it frequently represents a tendency to more frequent usage within one ethnic group rather than an either/or situation.

Even though colloquial forms differ from the standard ones primarily in the lexicon, there are also certain differing colloquial grammatical features which will be discussed in the appropriate section. Both lexical and grammatical features of colloquial forms can be used to determine the territory and, in some instances, ethnicity of the producer of the text.

1.4. Alphabets

Latin and Cyrillic scripts are used in contemporary Bosnian / Croatian / Serbian texts. ≍Serbs use both scripts (a recent survey shows that 40% of Serbs in Serbia use only the Latin script, 30% only the Cyrillic, while 30% uses both scripts, see *Politika*, 2003), while Croats and Bosniacs use only the Latin script. Although the Cyrillic script is officially the second script of both the Croatian and Bosniac standard, it is practically not used at all. Moreover, owing to the fact that the Cyrillic script is identified with the Serbian ethnic group, its usage is viewed unfavorably by most Croatian and Bosniac speakers. ≍

Given that the Latin script by and large shares characters with the English alphabet, an English learner should acquire it fully prior to proceeding with the Cyrillic script.

Chapter 1: General

Both Latin and Cyrillic scripts have thirty characters. Whereas they are all simple characters in the Cyrillic script, the Latin script consists of twenty-seven simple characters and three DIGRAPHS (i.e., sequences of two signs used to represent one sound), *lj*, *nj*, and *dž*. They are presented in the Table 1.4.1a in Latin alphabetical order with the Latin characters requiring the attention of an English speaker clearly marked. IPA (International Phonetic Alphabet) values are provided for each character.

<i>Latin</i>	<i>Cyrillic</i>	<i>IPA</i>	<i>name</i>	<i>English spelling approximates</i>
A a	А а	[a]	a	a as in <i>arm, harm</i>
B b	Б б	[b]	be	b as in <i>bad, body</i> pronounced in a less tense manner
C c	Ц ц	[ts]	ce	ts combination
Č č	Ч ч	[tʃ]	č	ch or tch as in <i>church, fetch</i> , pronounced harder than the English sound and harder than the ć below
Ć ć	Ћ ћ	[t]	ć	ch or tch as in <i>church, fetch</i> , pronounced softer than the English sound and softer than the č above
D d	Д д	[d]	de	d as in <i>do, hard</i> pronounced in a less tense manner
Dž dž	Џ џ	[dʒ]	dž	j as in <i>jeep, jungle</i> , pronounced harder than the English sound, and harder than the đ below
Đ đ	Ђ ј	[d]	đ	j as in <i>jeep, jungle</i> , pronounced softer than the English sound, and softer than the dž above
E e	Е е	[e]	e	e as in <i>red, bet</i>
F f	Ф ф	[f]	ef	f as in <i>fine, golf</i>
G g	Г г	[g]	ge	g as in <i>go, long</i>
H h	Х х	[h]	ha	h as in <i>head, hate</i>
I i	И и	[i]	i	i or ee as in <i>if, peel</i>
J j	Ј ј	[j]	j	y as in <i>yet, yolk</i>
K k	К к	[k]	ka	c as in <i>car, cast</i>
L l	Л л	[l]	el	l as in <i>lip, lore</i>
Lj lj	Љ љ	[ʎ]	lj	ly sequence
M m	М м	[m]	em	m as in <i>mother, me</i>
N n	Н н	[n]	en	n as in <i>next, north</i>
Nj nj	Њ њ	[ɲ]	nj	ny sequence
O o	О о	[o]	o	o as in <i>ox, ore</i>
P p	П п	[p]	pe	p as in <i>part, peace</i> pronounced in a less tense manner
R r	Р р	[r]	er	r as in <i>rare, rob</i>
S s	С с	[s]	es	s as in <i>seldom, yes</i>
Š š	Ш ш	[ʃ]	š	sh as in <i>shine</i> , pronounced harder than the English sound